II. REMARKS

Formal Matters

Claims 42-82 are pending after entry of the amendments set forth herein.

Claims 42-82 were examined. Claims 42-49 and 53-57 were rejected. Claims 50-52 and 58-82 were withdrawn from consideration.

Claims 42, 46 and 49 are amended. The amendments to the claims were made solely in the interest of expediting prosecution, and are not to be construed as an acquiescence to any objection or rejection of any claim. Support for the amendments to claims 42, 46, and 49 is found in the claims as originally filed, and throughout the specification, in particular at the following exemplary locations: Specification pages 8, 15, 19, and newly renumbered Figure 6B lanes 2 and 3. Accordingly, no new matter is added by these amendments.

New claims 83 – 100 have been added. Support for new claims 83 – 100 is found in the claims as originally filed, and throughout the specification, in particular at the following exemplary locations: Claims 83 – 91: Specification pages 8, 16, 18, 19, 20, and newly renumbered Figure 6B lane 1; Claims 92-100: Specification pages 8, 15, 17, and 21. Accordingly, no new matter is added by these new claims.

The disclosure has been amended in the specification to address objections noted in the Office Action. The specification has been amended on pages 7 and 10. Support for the amended material can be found in the claims as originally filed, and throughout the specification, in particular at the following exemplary locations: page 7: original claim 1; page 10: original claim 17.

The drawings have been amended to address objections noted in the Office Action. Original Figure 5B and original Figure 6B have been deleted. Original Figures 3C, 3D, 5C and 6C have been amended with revised figure numbers with replacement drawing sheets submitted herewith. The specification has also been amended to remove reference to the deleted figures.

Applicants respectfully request reconsideration of the application in view of the remarks made herein.

Examiner Interview

The undersigned Applicants' representative thanks Examiner Duffy for the courtesy of an in-person interview which took place on November 24, 2003, and which was attended by Examiner Duffy and Applicants' representatives Paula A. Borden and Edward J. Baba.

During the interview, the rejection of claims 42-49 and 53-57 under 35 U.S.C. §112, §102, and §103, was discussed. The amendments to the claims reflect the discussions, which took place during the interview.

Drawings

The drawings filed on 08/02/1999 have been objected to. Replacement figures have been submitted herewith, which replacement figures comply with the requirements for formal drawings. Withdrawal of these objections is respectfully requested.

Original Figures 3C and 3D have been amended to comply with the rules. Original Figure 3C has been amended to revise the figure numbers to Figures 3C-3Y. Accordingly, original Figure 3D has been amended to Figure 3Y, to reflect the renumbering of original Figure 3C.

Original Figures 5B and 6B have been deleted. Original Figures 5C and 6C have been amended to revise the figure numbers to reflect the deletion of Figures 5B and 6B. Replacement drawing sheets are provided herewith. The specification has also been amended to remove reference to the deleted figures.

Specification Objections

The disclosure was objected to because the text references claim numbers. The specification has been amended to remove the reference to claim numbers and alternatively, incorporate the language from the referenced claims into the text of the disclosure. The inserted material corresponds exactly to the text of the original claims that were referenced. Accordingly, no new matter has been added. Therefore, the Examiner is respectfully requested to withdraw the objection.

Claim objections

Claims 47, 48, and 49 were objected to under 37 C.F.R. 1.75(c) as allegedly being in improper dependent form for failing to further limit the subject matter of a previous claim.

Without conceding as to the correctness of this rejection, claims 47 and 48 have been canceled rendering the objection of these claims moot. In addition, claim 49 has been amended to remove reference to specific fragments of gp190/MSP1. Therefore, the Examiner is respectfully requested to withdraw this objection.

Rejection under 35 U.S.C.§112, first paragraph

New Matter

Claims 42-49 and 53-57 were rejected under 35 U.S.C. §112, first paragraph, as allegedly containing subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventors, at the time the application was filed, had possession of the claimed invention.

Specifically, the Office Action stated that the claims read on a reduction of any AT content in any MSP1 nucleotide sequence as compared to any other naturally occurring sequence. As suggested by the Examiner, independent claim 42 has been amended to recite "corresponding naturally occurring" in order to provide a reference point for comparison of a reduced AT content.

The Office Action also states that claims 47 and 48 stand rejected because the claims are drawn to a method of producing a complete gp190/MSP1 polypeptide wherein the nucleotide sequence further comprises an attachment signal or further comprises a signal peptide. Without conceding as to the correctness of this rejection, claims 47 and 48 have been canceled and new claims 83 to 100 have been added in their place. New claims 83 to 100 are directed to two other variations of the gp190/MSP1 protein where either the gp190/MSP1 lacks an attachment signal (claims 83-91) or it lack both an attachment signal and signal peptide (claims 92-100). Support for new claims 83 – 100 can be found in the claims as originally filed, and throughout the specification, in particular at the following exemplary

locations: Claims 83 – 91: Specification pages 8, 16, 18, 19, 20, and Figure 6C lane 1 (renumbered as Figure 6B); Claims 92-100: Specification pages 8, 15, 17, and 21. A schematic representation of examples of polypeptides recited in the claims is provided in the attached Exhibit 1, which was also made available to the examiner prior to the telephone interview on November 24, 2003. Accordingly, no new matter is added by these new claims.

Applicants submit that the new matter rejection of claims 42-49 and 53-57 under 35 U.S.C. §112, first paragraph, has been adequately addressed in view of the remarks set forth above. The Examiner is thus respectfully requested to withdraw the rejection.

Written Description

Claims 42-49 and 53-57 were rejected under 35 U.S.C. §112, first paragraph, as allegedly lacking written description in the specification. Specifically, the Examiner states that the specification "fails to describe the complete nucleotide sequences encoding naturally occurring gp190/MSP1 proteins corresponding to a representative number of these species, sufficient to describe the genus of nucleotides sequences that are modified to produce nucleotide sequences that are 'reduced' in their adenine-thymine content." In view of the remarks made below, applicants respectfully traverse this rejection.

Under MPEP § 2163.02, the standard for determining compliance with the Written Description requirement is whether the "specification conveys with reasonable clarity to those skilled in the art that, as of the filing date sought, applicant was in possession of the invention as now claimed." See, e.g., Vas-Cath, Inc. v. Mahurkar, 935 F.2d 1555, 1563-64, 19 USPQ2d 1111, 1117 (Fed. Cir. 1991). Essentially, the specification must "clearly allow persons of ordinary skill in the art to recognize that they invented what is claimed." See Vas-Cath, 935 F.2d at 1116.

In rejecting the claims the Office Action states that a representative number of nucleic acid sequences have not been provided for other Plasmodium species, therefore a skilled artisan could not "envision the detailed chemical structure of the encompassed nucleotide sequences that are used to produce the undescribed proteins of other at least 100 Plasmodium species and therefore conception is not achieved until reduction to practice has occurred...the nucleic acid itself is required." During the November 24, 2003 telephone interview, the Examiner indicated that if the nucleic acid sequence for

MSP-1 of a representative number of species of Plasmodium were known at the time the instant application was filed, such would be sufficient to overcome the written description rejection.

The Applicants submit that the nucleic acid sequences for the MSP-1 protein of a representative number of species of Plasmodium were known at the time the present application was filed. For example, Chang et al., Exp. Parisatol. 67(1):1-11 (1988) (Exhibit 2) discloses the nucleic acid encoding MSP-1 of *Plasmodium Falciparum* (Uganda-Palo Alto strain); Lewis et al., Mol. Biochem. Parisatol. 36(3):271-282 (1989) (Exhibit 3) discloses the nucleic acid encoding MSP-1 of *Plasmodium Yoelii*; Deleersnijder et al., Mol. Biochem. Parisatol. 43(2):231-244 (1990) (Exhibit 4) discloses the nucleic acid encoding MSP-1 of *Plasmodium Chaubaudi*; Del Portillo et al., Proc. Natl. Acad. Sci. 88:4030-4034 (1991) (Exhibit 5) discloses the nucleic acid encoding MSP-1 of *Plasmodium Vivax* (Belum strain); and Gibson et al., Mol. Biochem. Parisatol. 50(2):325-333 (1992) (Exhibit 6) closes the nucleic acid encoding MSP-1 of *Plasmodium Vivax* (Sal-1 strain). The Applicants note that the term Merzoite Surface Protein 1 (MSP1 or MSP-1) is also referred to in the literature as: Merzoite Surface Antigen 1 (MSA1 or MSA-1); Plasmodium Major Merzoite Surface Antigen (PMMSA); and Major Merzoite Surface Protein Precursor.

Since the sequence of the of MSP-1 gene of various species of Plasmodium were available, the methods disclosed in the present application could be readily applied by one skilled in the relevant art of molecular biology to these other MSP-1 genes to produce nucleotide sequences that are "reduced" in their adenine-thymine content. Accordingly, the Applicants submit that the written description rejection of claims 42-49 and 53-57 under 35 U.S.C. §112, first paragraph, has been adequately addressed in view of the remarks set forth above. The Examiner is thus respectfully requested to withdraw the rejection.

Enablement

Claim 53 was rejected under 35 U.S.C. §112, first paragraph, as allegedly containing subject matter which was not described in the specification in such a way as to enable one skilled in the art to which it pertains, or with which it is most nearly connected, to make and/or use in the invention. Without conceding to the correctness of this rejection, claim 53 has been canceled in the spirit of expediting prosecution. Accordingly, this rejection is rendered moot and the Examiner is thus respectfully requested to withdraw the rejection

Rejection under 35 U.S.C.§112, second paragraph

Claims 42-49 and 53-57 were rejected under 35 U.S.C. §112, second paragraph, as allegedly indefinite.

The Office Action stated that claim 42 presents a comparison between two sequences that are not structurally defined and therefore, the metes and bounds of the claim cannot be ascertained. Specifically, the Office Action notes that the term "complete" and the terms "naturally occurring nucleotide sequence" do not help to structurally define the two sequences. In the spirit of expediting prosecution, and without conceding as to the correctness of this rejection, claim 42 has been amended to remove the term "complete" and in its place add "having an approximate weight if 190 kD" in order to describe the gp190/MSP1 protein. In addition and the term "corresponding" has been added to define the naturally occurring sequence. These amendments were discussed during the November 24, 2003 telephone interview. Support for the amendments of claim 42 can be found in the claims as originally filed, and throughout the specification, in particular at the following exemplary locations: pages 8, 15, 19, and Figure 6C lanes 2 and 3 (renumbered as Figure 6B). The amendments of claim 42 have also been incorporated in newly presented independent claims 83 and 92.

Applicants submit that the rejection of claims 42-49 and 53-57 under 35 U.S.C. §112, second paragraph, has been adequately addressed in view of the remarks set forth above. The Examiner is thus respectfully requested to withdraw the rejection.

Rejection under 35 U.S.C.§102/103

Claims 42-49 and 53-57 were rejected under 35 U.S.C. §102(b) or §103 as allegedly unpatentable over Holder et al. ((1985) *Nature* 317:270-273; hereinafter "Holder").

The Office Action maintained the rejection from the Office Action dated November 20, 2001, and stated that Holder et al. teaches the production of specific fragments of the full length gp190/MSP1 from *P. falciparum*. The Examiner notes that the definition of the term "complete" on page 6 of the specification is inclusive of shorter forms, and that the claim does not define the specific sequence for comparison of the naturally occurring sequence.

As noted above, Claim 42 has been amended to remove the term "complete" and in its place add "having an approximate weight if 190 kD" in order to describe the gp190/MSP1 protein and the term "corresponding" has been added to define the naturally occurring sequence.

Holder does not render claims 42-49 and 53-57 obvious, as there is no mention in Holder of a method for producing gp190/MSP1 having an approximate molecular weight of 190 kD, much less a method of producing gp190/MSP1, comprising expressing a nucleotide sequence encoding gp190/MSP1 in a single expression vector. As stated in the specification, until the instant invention, there was not any successful cloning of the coding region for gp190/MSP1 having an approximate molecular weight of 190 kD. Holder does not disclose a method for solving this problem, nor does Holder suggest any such method. Accordingly, Holder cannot render the instant method as claimed obvious.

Accordingly, the Applicants submit that the rejection of claims 42-49 and 53-57 under 35 U.S.C.§102(b) or 103 has been adequately addressed in view of the amendments to the claims and remarks set forth above. Therefore, the Examiner is respectfully requested to withdraw the rejection and allow the application to proceed to issue.

III. CONCLUSION

Applicants submit that all of the claims are in condition for allowance, which action is requested. If the Examiner finds that a telephone conference would expedite the prosecution of this application, the Examiner is invited to telephone the undersigned at the number provided.

The Commissioner is hereby authorized to charge any underpayment of fees associated with this communication, including any necessary fees for extensions of time, or credit any overpayment to Deposit Account No. 50-0815, order number GRUE003.

Respectfully submitted,
BOZICEVIC, FIELD & FRANCIS LLP

Date: December 23, 2003

By: _____
Edward J. Baba
Registration No. 52,581

BOZICEVIC, FIELD & FRANCIS LLP
200 Middlefield Road, Suite 200
Menlo Park, CA 94025
Telephone: (650) 327-3400
Facsimile: (650) 327-3231

F:\DOCUMENT\GRUE (Gruenecker, Kinkeldey...)\003\resp OA 8-27-03 final.doc

U.S. App. No.: 09/269,874

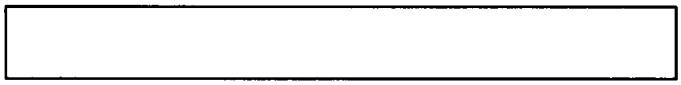
Group Art Unit: 1641

Eaminer: P.A. Duffy

Title: Recombinant Process for Preparing A Complete Malaria Antigen, GP190/MSP1

Protein: gp190/MSP1 – amino acids 1-1639 of SEQ ID NO:3
Gene: gp190s
Description: The complete protein with the attachment signal and signal peptide
Support: Specification pages 8 and 15

Protein: gp190/MSP1 – amino acids 1-1621 of SEQ ID NO:3
Gene: $gp190^{s1}$
Description: The protein with the signal peptide but lacking an attachment signal
Support: Specification pages 8 and 15



Protein: gp190/MSP1 – amino acids 20-1621 of SEQ ID NO:3
Gene: $gp190^{s2}$
Description: The protein lacking the attachment signal and signal peptide
Support: Specification pages 8 and 15

EXPERIMENTAL PARASITOLOGY 67, 1–11 (1988)

Plasmodium falciparum: Gene Structure and Hydropathy Profile of the Major Merozoite Surface Antigen (gp195) of the Uganda–Palo Alto Isolate

SANDRA P. CHANG, KENTON J. KRAMER, KAREN M. YAMAGA, ANN KATO, STEPHEN E. CASE, AND WASIM A. SIDDIQUI

Department of Tropical Medicine and Medical Microbiology, University of Hawaii School of Medicine, Honolulu, Hawaii, 96816, U.S.A.

(Accepted for publication 19 April 1988)

CHANG, S. P., KRAMER, K. J., YAMAGA, K. M., KATO, A., CASE, S. E., AND SIDDIQUI, W. A. 1988. *Plasmodium falciparum:* Gene structure and hydropathy profile of the major merozoite surface antigen (gp195) of the Uganda–Palo Alto Isolate. *Experimental Parasitology* 67, 1–11. The gene encoding the 195,000-Da major merozoite surface antigen (gp195) of the FUP (Uganda-Palo Alto) isolate of *Plasmodium falciparum*, a strain widely used for monkey vaccination experiments, has been cloned and sequenced. The translated amino acid sequence of the FUP gp195 protein is closely related to the sequences of corresponding proteins of the CAMP (Malaysia) and MAD-20 (Papua New Guinea) isolates and more distantly related to those of the Wellcome (West Africa) and K1 (Thailand) isolates, supporting the proposed allelic dimorphism of gp195 within the parasite population. The prevalence of dimorphic sequences within the gp195 protein suggests that many gp195 epitopes would be group-specific. Despite the extensive differences in amino acid sequence between gp195 proteins of these two groups, the hydropathy profiles of proteins representative of both groups are very similar. The conservation of overall secondary structure shown by the hydropathy profile comparison indicates that gp195 proteins of the various *P. falciparum* isolates are functionally equivalent. This information on the primary structure of the FUP gp195 protein will enable us to evaluate the possible roles of conserved, group-specific and variable epitopes in immunity to the blood stage of the malaria parasite.

INDEX DESCRIPTORS AND ABBREVIATIONS: *Plasmodium falciparum;* Protozoa, parasitic; Malaria, human; Merozoite; Vaccine; Major merozoite surface antigen (gp195); Uganda–Palo Alto isolate (FUP); Dalton (Da); Kilodalton (kDa); Kilobase (kb); Deoxyribonucleic acid (DNA); Base pair (bp).

INTRODUCTION

Plasmodium falciparum is the causative agent of the most serious form of human malaria. The surface of the malaria parasite undergoes drastic antigenic changes during its complex life cycle. The predominant surface antigen of the sporozoite produced during the sexual cycle of the parasite in the *Anopheles* mosquito host is the circumsporozoite protein (Nussenzweig and Nussenzweig 1985). After injection of the sporozoite into the bloodstream of the vertebrate host and its uptake into hepatocytes the circumsporozoite protein is lost (Danforth *et al.* 1978). Little is known about surface antigens of the parasite during the hepatic stage of asexual development. However, surface antigens of the erythrocytic stages of the parasite life cycle have been well studied. The major surface antigens of the *Plasmodium falciparum* merozoite, the erythrocytic invasive stage of the parasite, are derived from a precursor glycoprotein with a molecular weight of 185–195,000 (gp195) (Freeman and Holder 1983; Hall *et al.* 1983; Holder and Freeman 1984). The gp195 precursor protein is synthesized during the late erythrocytic stage of development and is proteolytically processed to

0014-4894/88 $3.00

MN 55802
countries
countries

lower molecular weight fragments (Freeman and Holder 1983; Hall *et al.* 1984a). Three of these fragments of 83,000, 42,000, and 19,000 Da have been detected on the merozoite surface (Holder and Freeman 1984). An approximately 80,000-Da processing fragment of the gp195 has been localized in the surface coat of the merozoite by immunoelectron microscopy (Heidrich *et al.* 1986).

The genes encoding the gp195 protein of several parasite isolates have been cloned and sequenced, revealing sequence polymorphisms among different genes (Hall *et al.* 1984b; Holder *et al.* 1985; Mackay *et al.* 1985; Weber *et al.* 1986; Tanabe *et al.* 1987; Peterson *et al.* 1988). Gp195-related polypeptides are candidates for a blood stage human malaria vaccine. Results of three monkey vaccination experiments using gp195-derived immunogens showed partial (Hall *et al.* 1984b; Perrin *et al.* 1984) to complete protection (Siddiqui *et al.* 1987). In each experiment, challenge has been with the *P. falciparum* FUP isolate, but only in the latter experiment were the monkeys immunized with gp195 purified from FUP parasites. In order to develop a fully protective recombinant polypeptide or synthetic peptide vaccine based on gp195 and to evaluate the significance of antigenic polymorphism of this protein in protective immunity, it appears crucial to define the structure of the FUP gp195 gene.

This study presents the DNA sequence of the FUP gp195 gene and compares its translated amino acid sequence to others that have been published. In addition, the amino acid sequence dimorphism of gp195 proteins is correlated to secondary structure as predicted by hydropathy analysis.

Materials and Methods

Parasites. The Uganda–Palo Alto (FUP) strain of *Plasmodium falciparum* was originally isolated from a patient who had contracted the infection in Uganda and was hospitalized at Stanford Medical Center, Palo Alto, California, in 1966. In 1967, blood-induced infections with this isolate were established in *Aotus trivirgatus* monkeys at Stanford (Geiman and Meagher 1967). The FUP strain was maintained by serial passage in *Aotus* monkeys by Dr. Schmidt of the Southern Research Institute, Birmingham, Alabama. In 1970, the FUP monkey-passaged strain was obtained from Dr. Schmidt and maintained in this laboratory at the University of Hawaii by serial passage in *Aotus* monkeys. In 1977, continuous *in vitro* cultures of the FUP strain in human erythrocytes were established at the University of Hawaii and have been maintained in this laboratory since that time. The FUP parasites used in this study were derived from *in vitro* cultures.

Isolation of P. falciparum DNA. DNA was isolated from cultured FUP strain *P. falciparum* using the Trager and Jensen (1976) culture technique with modifications (Siddiqui and Palmer 1981) and standard DNA extraction methods (Maniatis *et al.* 1982). The Protoclone bacteriophage λ gt10 system (Promega Biotec, Madison, WI, U.S.A.) was used to generate an FUP *P. falciparum* genomic library. FUP DNA (0.3 μg, 0.1 pmole) was digested with 3 units of the restriction endonuclease *Eco*RI (Boehringer Mannheim, Indianapolis, IN, U.S.A.) and ligated to 0.5 μg (0.17 pmole) *Eco*RI-digested λ gt10 DNA (Promega Biotec) with 1 unit T4 polynucleotide ligase (Promega Biotec). The resultant recombinant phage were grown in *Escherichia coli* strain C600ΔHFL cells, and generated a library of 2.5×10^6 plaque-forming units in which 87% of the phage contained inserts.

Preparation of synthetic oligonucleotides. Oligonucleotides used as hybridization probes were synthesized as pairs of 30-mers overlapping by 10 base pairs. Probes were synthesized using methoxy phosphoramidites or β-cyanoethyl phosphoramidites on Applied Biosystems DNA synthesizers (Foster City, CA, U.S.A.) according to the manufacturers recommendations. The oligonucleotides were radiolabeled by fill-in reactions using high specific activity (3000 Ci/mmole) 5′-[α-^{32}P]deoxycytidine triphosphates and -deoxyadenosine triphosphates (Amersham, Arlington Heights, IL, U.S.A.) and Klenow DNA polymerase I (Boehringer Mannheim). Hybridization to phage DNA on nitrocellulose filters and washing were carried out as described by Ullrich *et al.* (1984).

Restriction map analysis and DNA sequencing. λ gt10 recombinant phage inserts were subcloned into pUC plasmids for restriction mapping and into M13mp18 or M13mp19 phage for DNA sequencing. M13 subclones were sequenced by the enzymatic method (Sanger *et al.* 1977) using a universal M13 primer (Pharmacia, Piscataway, NJ, U.S.A.) or specific primers complementary to the insert sequence. Sequencing primers were synthesized using β-cyanoethyl phosphoramidites as described above for oligonucleotide probes. DNA sequence data were analyzed

P. j

using computer resources
sponsored BIONET Natior
Molecular Biology. Protein
ally analyzed using the con
by Pauletti *et al.* (1985) and

Resu

A λ gt10 library of
DNA was screened w
oligonucleotide probe
served sequences in th
Wellcome and K1 st
1985; Mackay *et al.* 1
(10-1, 3-1, and 18-1)
coding region of the st
(Fig. 1) and were map
endonucleases and sub
quencing vectors. The
these three λ clones w

The complete DNA
acid translation of the
gene is presented in Fi
molecular weight of t
196,245. There are
glycosylation sites (Sn
tein contains 20 cystei
teines are conserved
lates and 13 are located
the hydrophobic carbo

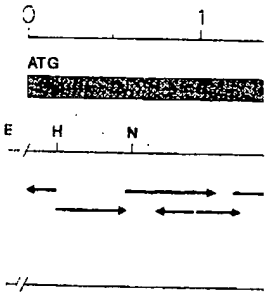


Fig. 1. Map of the gp
box shows the positions
map indicates the restri
subclones of the 18-1 λ g
indicate the sequencing s
primers. Restriction end

blood-induced infec-
blished in *Aotus tri-*
eiman and Meagher
ained by serial pas-
midt of the Southern
Alabama. In 1970,
was obtained from
is laboratory at the
sage in *Aotus* mon-
cultures of the FUP
e established at the
n maintained in this
P parasites used in
tro cultures.

DNA was isolated
ciparum using the
chnique with mod-
981) and standard
s *et al.* 1982). The
item (Promega Bio-
ed to generate an
y. FUP DNA (0.3
nits of the restric-
er Mannheim, In-
d to 0.5 μg (0.17
(Promega Biotec)
(Promega Biotec).
re grown in *Esch-*
and generated a
nits in which 87%

eotides. Oligonu-
es were synthe-
by 10 base pairs.
oxy phosphora-
dites on Applied
oster City, CA,
rs recommenda-
labeled by fill-in
(3000 Ci/mmole)
es and -deoxy-
am, Arlington
A polymerase I
n to phage DNA
were carried out

sequencing. λ
subcloned into
ping and into
A sequencing.
the enzymatic
universal M13
U.S.A.) or spe-
sert sequence.
using β-cyano-
bove for oligo-
were analyzed

using computer resources provided by the N.I.H. sponsored BIONET National Computer Resource for Molecular Biology. Protein sequences were additionally analyzed using the computer programs described by Pauletti *et al.* (1985) and Gotoh (1986).

Results

A λ gt10 library of FUP strain genomic DNA was screened with several synthetic oligonucleotide probes based on the conserved sequences in the gp195 genes of the Wellcome and K1 strains (Holder *et al.* 1985; Mackay *et al.* 1985). Three λ clones (10-1, 3-1, and 18-1) contained the entire coding region of the strain FUP gp195 gene (Fig. 1) and were mapped using restriction endonucleases and subcloned into M13 sequencing vectors. The M13 subclones of these three λ clones were sequenced.

The complete DNA sequence and amino acid translation of the strain FUP gp195 gene is presented in Fig. 2. The calculated molecular weight of the entire protein is 196,245. There are 15 potential *N*-glycosylation sites (Snider 1984). The protein contains 20 cysteine residues; 19 cysteines are conserved among various isolates and 13 are located immediately before the hydrophobic carboxy terminal region.

The amino acid translation of the FUP gp195 gene has been aligned with the sequences of other isolates (Fig. 3) using the computer algorithm described by Gotoh (1986). Based on this alignment the gp195 sequence can be divided into three types of regions. The first type is the conserved region (85–100% sequence identity). Conserved regions are located at the amino terminus and the carboxy terminus, as well as at internal segments of the protein. The second type of region is the variable repeat region which differs greatly among isolates in both sequence and length and is located toward the amino terminus. The third type of region is designated group-specific because it appears to exist in two forms which differ greatly in sequence (42–46% sequence identity). Over one-half of the gp195 protein consists of group-specific sequences. Others have previously recognized this pattern of polymorphism of the gp195 gene in the parasite population and refer to it as allelic dimorphism (Tanabe *et al.* 1987). The FUP protein belongs to the same dimorphic group as the gp195 of the Papua New Guinea isolate MAD20, differing primarily at the variable repeat region.

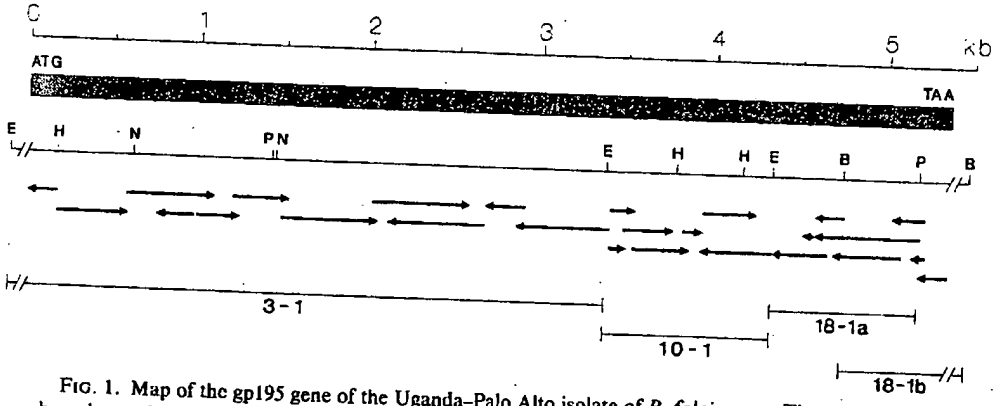


FIG. 1. Map of the gp195 gene of the Uganda–Palo Alto isolate of *P. falciparum*. The upper striped box shows the positions of the start and stop codons and the length of the coding region. The lower map indicates the restriction endonuclease sites of the λ gt10 clones (3-1 and 10-1) and the M13 subclones of the 18-1 λ gt10 clone (18-1a and 18-1b) shown on the bottom of the figure. The arrows indicate the sequencing strategy using either an M13 universal primer or gp195-specific oligonucleotide primers. Restriction endonucleases: E, *Eco*RI; H, *Hind*III; N, *Nde*I; B, *Bgl*II; and P, *Pst*I.

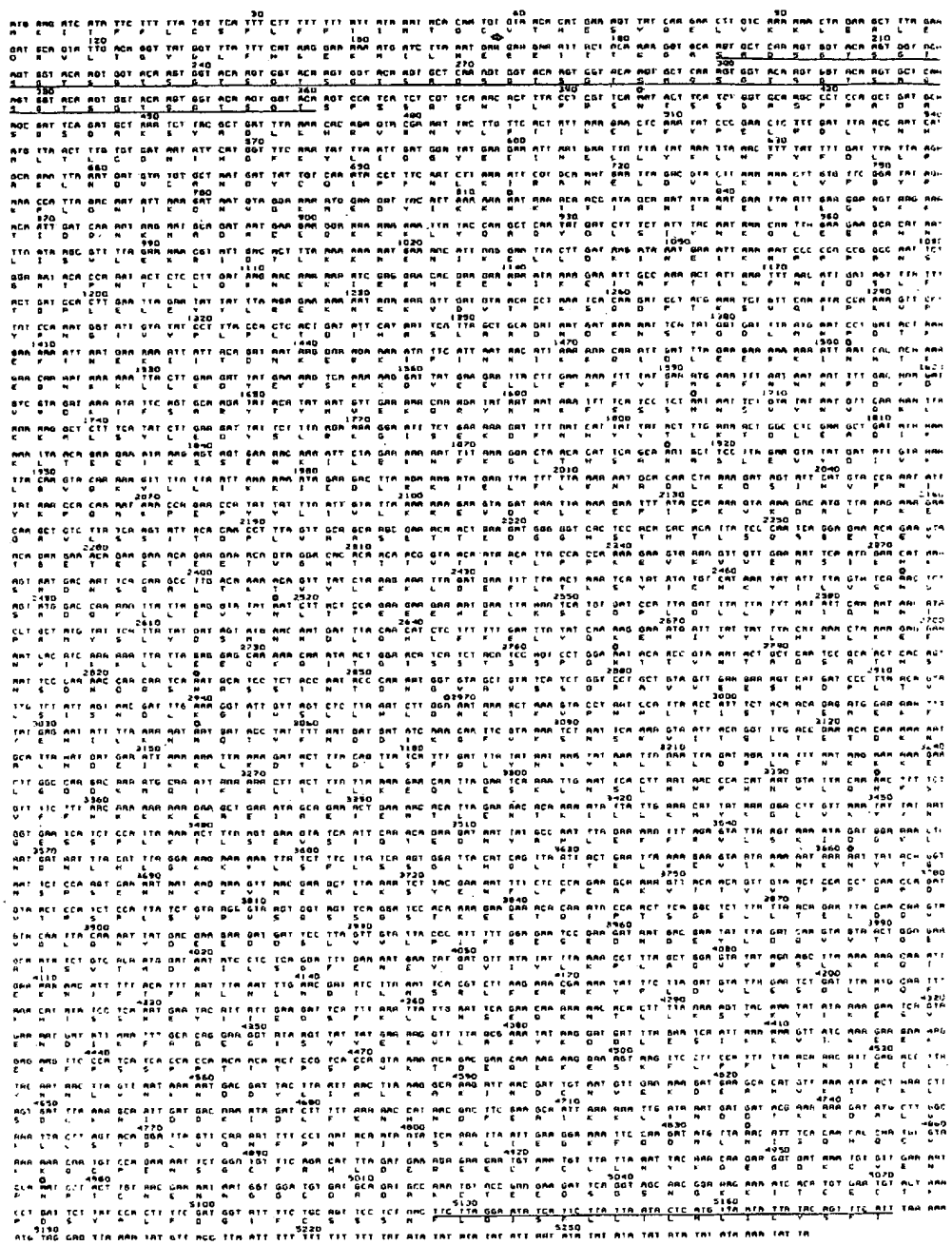
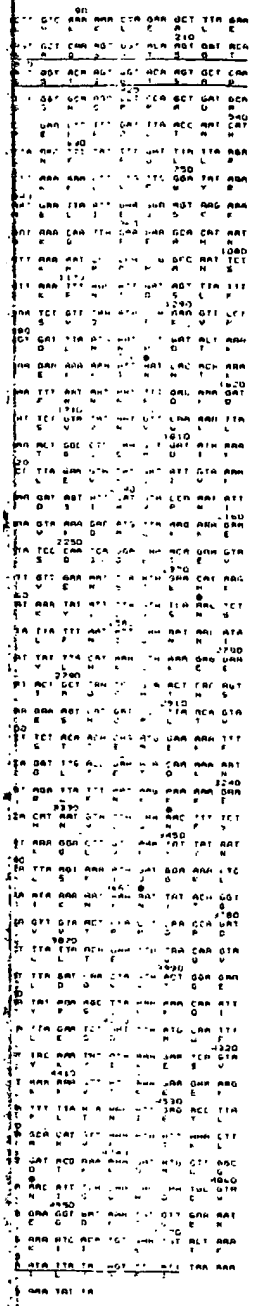


FIG. 2. Nucleotide sequence of the Uganda–Palo Alto (FUP) gp195 gene. The deduced amino acid sequence of the open reading frame is shown below the nucleotide sequence. The potential signal peptidase site is indicated with an arrowhead at position 57. The variable repeat region is underlined beginning at position 190 and the hydrophobic carboxy terminal sequence is overlined beginning at position 5125. Potential *N*-glycosylation sites are denoted by solid circles.

Several substitu
FUP amino aci
quence which it
sequence), alon
relative to MAL
eral single ami
tered throughou
the FUP and M
very similar ov
in contrast to th
lationship of F
The variable r
gp195 protein is
laysian CAMP
In fact, the pu
the CAMP gp19
identical to the
the FUP gene.

The extensiv
the group-spec
raise the possi
may represent
teins. Hydropa
secondary stru
interactions (K
been used to c
servation amor
(Simpson *et al.*
dropathy profil
dimorphic gp19
tion of the me
(Kyte and Do
1985) to determ
would be prec
(Fig. 4). As a
served sequen
dropathy patte
gions also diff
ever, regions
group-specific
similar hydrop
tle changes in
hydrophobicity
interactions dif
1000 (FUP) ar
correspond to
that are lackin
overall conser



deduced amino acid
The potential signal
region is underlined
derlined beginning at

Several substitutions can also be noted at FUP amino acid positions 344–376 (a sequence which it shares with the Wellcome sequence), along with a 6-base-pair deletion relative to MAD20 at position 782, and several single amino acid substitutions scattered throughout the gene. Nevertheless, the FUP and MAD20 gp195 proteins are very similar over most of their sequences, in contrast to the more distant sequence relationship of FUP and Wellcome gp195. The variable repeat region of the FUP gp195 protein is the same as that of the Malaysian CAMP isolate (Weber *et al.* 1986). In fact, the published partial sequence of the CAMP gp195 gene (Weber *et al.* 1986) is identical to the corresponding sequence of the FUP gene.

The extensive sequence differences of the group-specific regions of gp195 genes raise the possibility that these dimorphs may represent functionally divergent proteins. Hydropathy profiles, which reflect secondary structure as predicted by water interactions (Kyte and Doolittle 1982), have been used to demonstrate functional conservation among distantly related proteins (Simpson *et al.* 1987). We compared the hydropathy profiles of the FUP and Wellcome dimorphic gp195 proteins using an adaptation of the method of Kyte and Doolittle (Kyte and Doolittle 1982; Pauletti *et al.* 1985) to determine whether these proteins would be predicted to differ in structure (Fig. 4). As anticipated, regions of conserved sequence displayed identical hydropathy patterns. The different repeat regions also differed in hydropathy. However, regions containing dimorphic or group-specific sequences maintained very similar hydropathy patterns with only subtle changes in degree of hydrophilicity or hydrophobicity. A few regions where water interactions differed were at positions 900–1000 (FUP) and 1460–1640 (FUP), which correspond to insertions into the FUP gene that are lacking in the Wellcome gene. The overall conservation of the hydropathy pattern indicates that despite extensive amino acid differences the basic structure of the gp195 protein has been conserved among *Plasmodium falciparum* isolates.

Discussion

The determination of the complete DNA sequence of the FUP isolate gp195 gene extends our understanding of the degree of polymorphism of the major merozoite surface coat protein of *Plasmodium falciparum*. Its similarity to the previously reported gp195 sequence of the Papua New Guinea isolate (Tanabe *et al.* 1987) indicates that outside of the variable repeat region the polymorphism of this antigen is not extreme. This supports the proposal of others that the parasite population may be represented by two allelic groups, or dimorphs, which can undergo recombination during the sexual cycle in the mosquito vector to produce hybrid proteins, such as that observed for the Thailand K1 isolate (Tanabe *et al.* 1987). The identity between the partial sequence of the Malaysian CAMP isolate gp195 gene (Weber *et al.* 1986) and the corresponding region of the FUP gene is surprising since these isolates were obtained from distinct geographical areas. Nearly identical DNA sequences (8 nucleotide differences) were reported for the gp195 genes of the MAD20 and FC27 isolates (Tanabe *et al.* 1987; Peterson *et al.* 1988); however, these two isolates were both obtained from Papua New Guinea and thus may be derived from the same parasite population. The FUP and CAMP isolates were also found to be similar in sensitivity to several antimalarial drugs *in vitro* (Siddiqui *et al.* 1972) although they differed in drug sensitivity *in vivo* (Degowin and Powell, 1965). The genetic relationship between parasites of these two isolates is being investigated further using molecular probes specific for several genetic loci of *P. falciparum*.

Sequence comparisons between gp195 genes of different isolates indicate that

amino acid sequences of the dimorphic proteins can differ by greater than 50% in certain parts of the gene. Given such extensive sequence dissimilarities, we examined whether structural differences as reflected by hydropathy profiles could be observed. A comparison of hydropathy profiles of the FUP and Wellcome gp195 proteins, which represent the dimorphic groups, indicated that outside of the variable repeat regions the hydropathy patterns of the two proteins were very similar. These results suggest that the overall structure of the gp195 proteins of different *P. falciparum* isolates is conserved. Such structural conservation suggests that the group-specific regions of gp195 proteins share the same function. However, this function may be more dependent on the overall conformation of the protein than on its primary sequence. Alternatively, it is possible that group-specific regions may be involved in a function which can be carried out using two different pathways, such as the interaction of the parasite with different receptors on the host cell, as has been suggested by others (Tanabe *et al.* 1987).

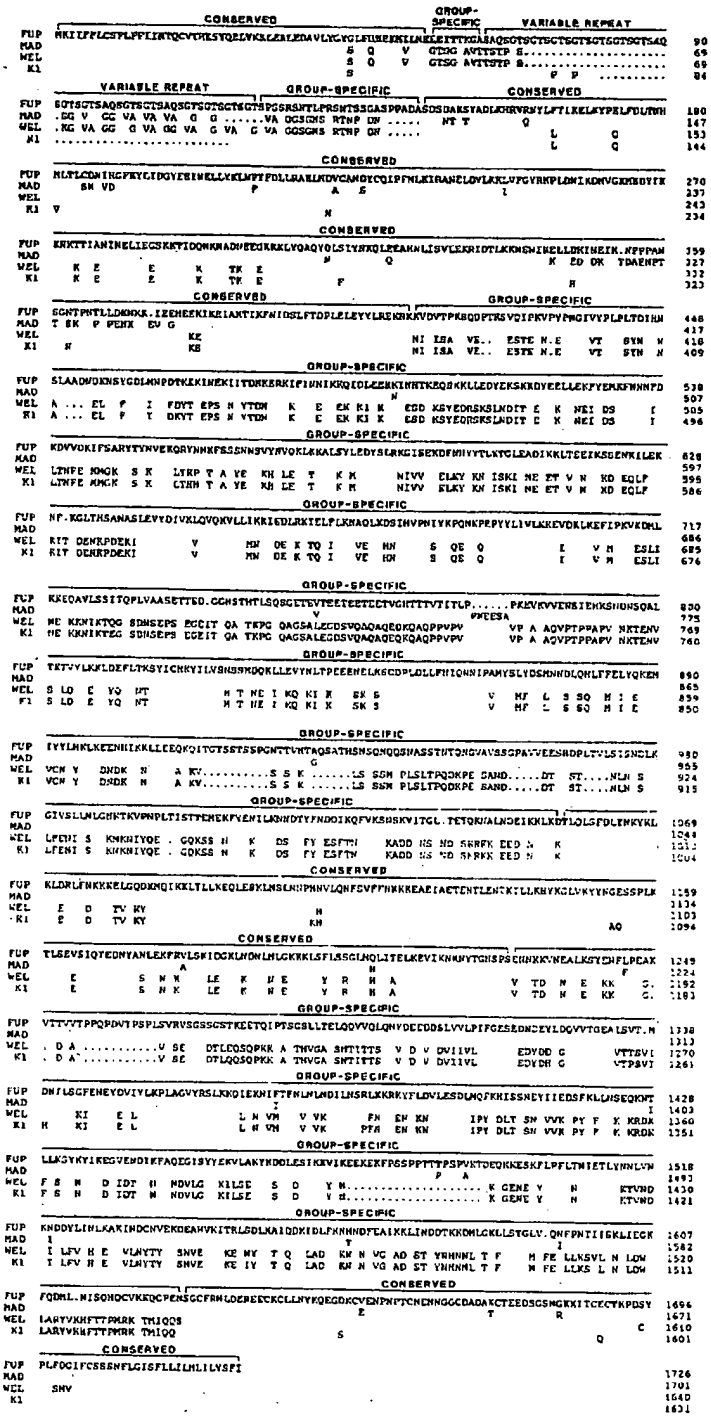
In designing an effective malaria vaccine, it is essential that the level of antigenic polymorphism of candidate antigens be carefully assessed. It is generally recognized that highly variable, isolate-specific sequences such as the variable repeats of gp195 are less attractive for vaccine development than conserved sequences. Several regions of the gp195 genes are highly conserved among all of the sequences that have been studied. These conserved regions are located on either side of the variable repeats at the amino terminal end of the protein, at the carboxy terminal region, and between the two large group-specific regions (Fig. 3). These conserved regions would be located on two to three distinct processing fragments on the mature merozoite surface (Lyon *et al.* 1986) and may be important in the merozoite invasion process. However, it remains to be shown whether epitopes contained in these regions are more relevant to immunity than those in other, less-conserved regions. Both conserved and group-specific regions of this protein have a generally hydrophilic character and exposure of these regions on the protein surface would allow them to be recognized as antigenic determinants (Hopp and Woods 1981). The extensive amino acid substitutions in group-specific regions of the gp195 protein make it likely that proteins of different groups would be antigenically distinct. While immunity developed against these regions would probably be group-specific, the evidence that there are a limited number (possibly only two forms) of group-specific regions make it feasible to consider including both of these regions in a recombinant vaccine.

Information on the primary structure of gp195 in the FUP strain enables us to reevaluate vaccination experiments in which monkeys immunized with antigen from another strain of *P. falciparum* were challenged with parasites of the FUP strain. While it must be recognized that these experiments differed in experimental detail, it is still informative to discuss them in light of this new information. Hall *et al.* (1984b) immunized *Saimiri* monkeys with monoclonal antibody-purified p190, the gp195-equivalent of the K1 (Thailand) strain of the parasite, and challenged these animals with the FUP strain. Two of three immunized

FIG. 3. Comparison of the amino acid sequences encoded by the gp195 gene of the FUP-Uganda, MAD20-Papua New Guinea (Tanabe *et al.* 1987), Wellcome-Lagos (Holder *et al.* 1985), and K1-Thailand (Mackay *et al.* 1985) strains of *P. falciparum*. Alignment was done using the Gotoh algorithm (Gotoh 1986). Shared sequences are indicated by blank spaces and gaps are indicated by periods. Conserved, variable repeat, and group-specific regions are indicated as the respectively designated, overlined sequences.

group-specific re-
onserved regions
o to three distinct
the mature mero-
1986) and may be
bite invasion pro-
ins to be shown
ed in these regions
unity than those in
gions. Both con-
ic regions of this
hydrophilic char-
ese regions on the
ow them to be rec-
terminants (Hopp
extensive amino
p-specific regions
it likely that pro-
would be antigen-
munity developed
ould probably be
ce that there are a
only two forms) of
ake it feasible to
these regions in a

mary structure of
enables us to re-
eriments in which
antigen from an-
arum were chal-
the FUP strain.
ed that these ex-
erimental detail, it
ss them in light of
et al. (1984b) im-
with monoclonal
90, the gp195-
land) strain of the
hese animals with
three immunized



le FUP-Uganda,
1985), and K1-
Gotoh algorithm
ated by periods.
vely designated,

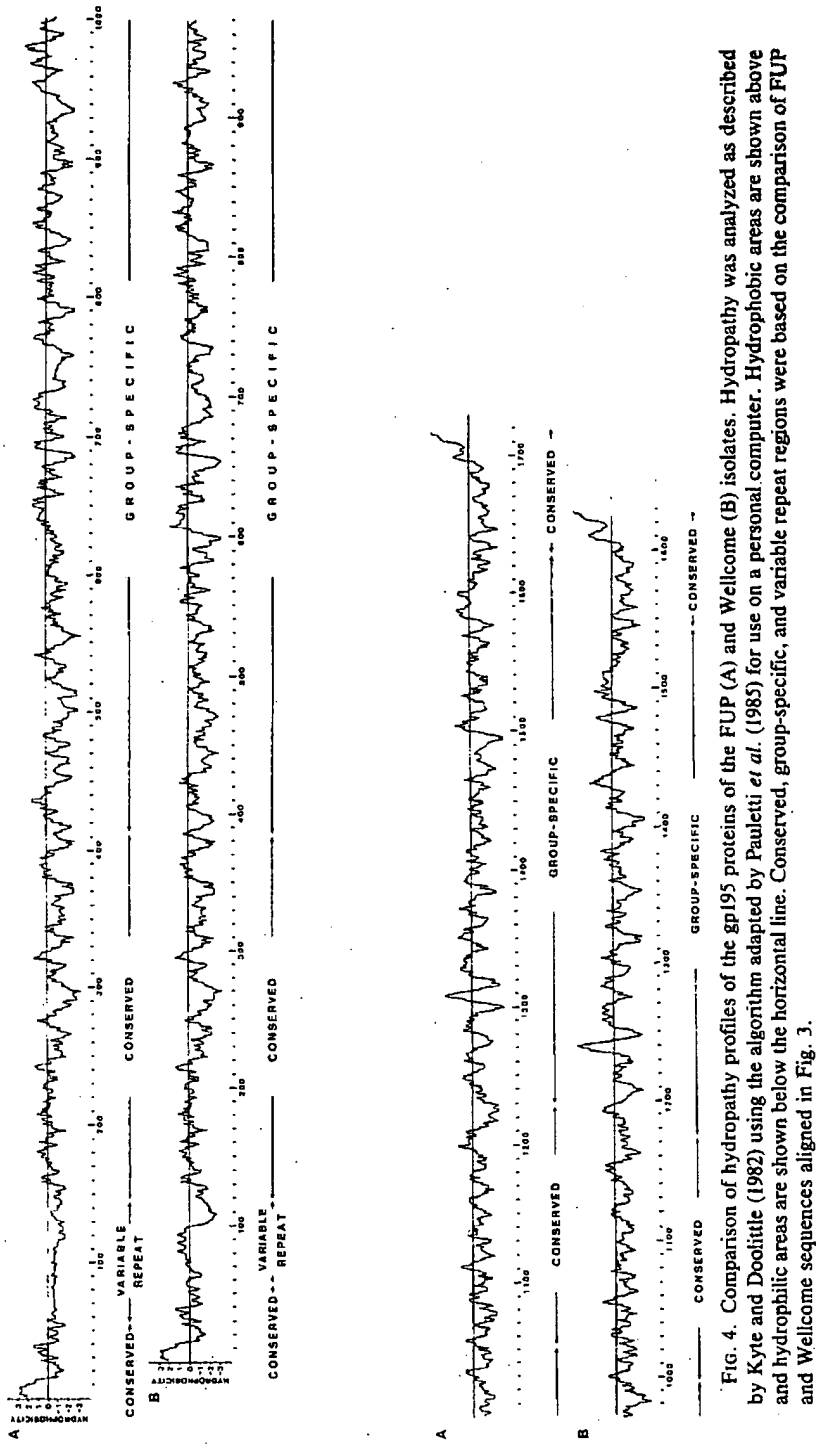


FIG. 4. Comparison of hydropathy profiles of the gp195 proteins of the FUP (A) and Wellcome (B) isolates. Hydropathy was analyzed as described by Kyte and Doolittle (1982) using the algorithm adapted by Pauletti *et al.* (1985) for use on a personal computer. Hydrophobic areas are shown above and hydrophilic areas are shown below the horizontal line. Conserved, group-specific, and variable repeat regions were based on the comparison of FUP and Wellcome sequences aligned in Fig. 3.

animals were partia
asite infection (5–10
third animal reactec
munized controls a
ment. Siddiqui *et*
(Vietnam) strain p
tion, also obtained
of immunized *Aotu*
with FUP strain p
contrast with the c
tained in *Aotus* mo
purified FUP gp19
parasites of the sam
1987). The results
FUP heterologous
are particularly inte
for a 36-bp deletion
first 375 amino acid
are 99.9% identical
Fig. 3). The repeat
FUP strain is mad
are also found in
[SAQ(SGT)n], the
the larger number
peats (n) and overa
FUP repeat region
third type of repea
beyond residue 375
sequence similarity
proteins drops bel
protection may hav
experiment by in
epitopes located wi
region of the polype
tion may require
served epitopes loc
and/or to confor
pressed by the lon
FUP polypeptide.
conserved epitope
ported by the fin
(1986), in which
nized with a conser
tide of gp195 were
tected from challer
asite. Most recentl
have shown that a
sponding to a mod

FIG. 4. Comparison of hydropathy profiles of the gp195 proteins of the FUP (A) and Wellcome (B) isolates. Hydropathy was analyzed as described by Kyte and Doolittle (1982) using the algorithm adapted by Pauletti *et al.* (1985) for use on a personal computer. Hydrophobic areas are shown above and hydrophilic areas are shown below the horizontal line. Conserved, group-specific, and variable repeat regions were based on the comparison of FUP and Wellcome sequences aligned in Fig. 3.

animals were partially protected from parasite infection (5–10% parasitemias) while a third animal reacted similarly to the unimmunized controls and required drug treatment. Siddiqui *et al.* (1978), using FVO (Vietnam) strain parasites for immunization, also obtained only partial protection of immunized *Aotus* monkeys challenged with FUP strain parasites. These results contrast with the complete protection obtained in *Aotus* monkeys immunized with purified FUP gp195 and challenged with parasites of the same strain (Siddiqui *et al.* 1987). The results of the strain K1/strain FUP heterologous vaccination experiment are particularly interesting because except for a 36-bp deletion in the repeat region, the first 375 amino acids of K1 and FUP gp195 are 99.9% identical (differ by 16 base pairs, Fig. 3). The repeat region of the challenge FUP strain is made up of repeats which are also found in the K1 polypeptide [SAQ(SGT)n], the only differences being the larger number of consecutive SGT repeats (n) and overall greater length of the FUP repeat region and the absence of a third type of repeat unit (SGP). However, beyond residue 375 of the K1 sequence, the sequence similarity between K1 and FUP proteins drops below 50%. While partial protection may have been achieved in this experiment by immunity to conserved epitopes located within the amino terminal region of the polypeptide, complete protection may require immunity to nonconserved epitopes located beyond this region and/or to conformational epitopes expressed by the longer repeat region of the FUP polypeptide. The involvement of nonconserved epitopes in immunity is supported by the findings of Cheung *et al.* (1986), in which *Saimiri* monkeys immunized with a conserved amino terminal peptide of gp195 were also incompletely protected from challenge with the malaria parasite. Most recently Patarroyo *et al.* (1987) have shown that a synthetic peptide corresponding to a moderately conserved region at the amino terminal end of gp195 (residues 43–53) contributed to the development of protective immunity but could not alone protect *Aotus* monkeys against malaria.

The mounting evidence linking the gp195 molecule to protection in the monkey model (Hall *et al.* 1984b; Perrin *et al.* 1984; Cheung *et al.* 1986; Patarroyo *et al.* 1987; Siddiqui *et al.* 1987), along with the growing number of characterized gp195 genes of different *P. falciparum* isolates, provides a powerful basis for development of a blood stage malaria vaccine. Information on the structural relatedness of the different gp195 genes permits a rational design of vaccination experiments which simulate the polymorphism encountered in nature. It also allows us to evaluate the possible need for a multivalent gp195 vaccine to achieve clinical immunity in a susceptible population.

ACKNOWLEDGMENTS

We are grateful to J. Baudler and A. Dasch and Drs. S. Horvath and L. Hood for providing oligonucleotide probes. We thank Drs. L. Tam, G. Hui, L. Simpson, and M. Rittenberg for critical reading of the manuscript. We acknowledge the technical advice of R. Lau and Dr. T. Humphreys during the early stages of this work. This research was supported by United States Agency for International Development Grant DPE-0453-A-00-4039 to W.A.S., University of Hawaii Research, and Training Revolving Fund Grant 21R87860F728B210 to S.P.C., and Bionet Grant 1 U41 RR-01685-03.

REFERENCES

CHEUNG, A., LEBAN, J., SHAW, A. R., MERKLI, B., STOCKER, J., CHIZZOLINI, C., SANDER, C., AND PERRIN, L. H. 1986. Immunization with synthetic peptides of a *Plasmodium falciparum* surface antigen induces antimerozoite antibodies. *Proceedings of the National Academy of Science USA* 83, 8328–8332.

DANFORTH, H. D., ORJIH, A. U., AND NUSSENZWEIG, R. S. 1978. Immunofluorescent staining of exoerythrocytic schizonts of *Plasmodium berghei* in fixed liver tissue with stage-specific immune serum. *Journal of Parasitology* 64, 1123–1125.

DEGOWIN, R. L. AND POWELL, R. D. 1965. Drug resistance of a strain of *Plasmodium falciparum* from Malaya. *American Journal of Tropical Medicine and Hygiene* 14, 519–528.

FREEMAN, R. R., AND HOLDER, A. A. 1983. Surface antigens of malaria merozoites. A high molecular weight precursor is processed to an 83,000 mol. wt. form expressed on the surface of *Plasmodium falciparum* merozoites. *Journal of Experimental Medicine* 158, 1647–1653.

GEIMAN, Q. M. AND MEAGHER, M. 1967. Susceptibility of a New World monkey to *Plasmodium falciparum* from man. *Nature* 215, 437–439.

GOTOH, O. 1986. Alignment of three biological sequences with an efficient traceback procedure. *Journal of Theoretical Biology* 121, 327–337.

HALL, R., MCBRIDE, J., MORGAN, G., TAIT, A., ZOLG, J. W., WALLIKER, D., AND SCAIFE, J. 1983. Antigens of the erythrocytic stages of the human malaria parasite *Plasmodium falciparum* detected by monoclonal antibodies. *Molecular and Biochemical Parasitology* 7, 247–265.

HALL, R., OSLAND, A., HYDE, J. E., SIMMONS, D. L., HOPE, I. A., AND SCAIFE, J. G. 1984a. Processing, polymorphism, and biological significance of P190, a major surface antigen of the erythrocytic forms of *Plasmodium falciparum*. *Molecular and Biochemical Parasitology* 11, 61–80.

HALL, R., HYDE, J. E., GOMAN, M., SIMMONS, D. L., HOPE, I. A., MACKAY, M., SCAIFE, J., MERKLI, B., RICHLE, R., AND STOCKER, J. 1984b. Major surface antigen gene of a human malaria parasite cloned and expressed in bacteria. *Nature* (*London*) 311, 379–382.

HEIDRICH, H.-G., MATZNER, M., MIETTINEN-BAUMANN, A., AND STRYCH, W. 1986. Immunoelectron microscopy shows that the 80,000-dalton antigen of *Plasmodium falciparum* merozoites is localized in the surface coat. *Zeitschift fur Parasitenkunde* 72, 681–683.

HOLDER, A. A., AND FREEMAN, R. R. 1984. The three major surface antigens on the surface of *Plasmodium falciparum* merozoites are derived from a single high molecular weight precursor. *Journal of Experimental Medicine* 160, 624–629.

HOLDER, A. A., LOCKYER, M. J., ODINK, K. G., SANDHU, J. S., RIVEROS-MORENO, V., NICHOLLS, S. C., HILLMAN, Y., DAVEY, L. S., TIZARD, M. L. V., SCHWARZ, R. T., AND FREEMAN, R. R. 1985. Primary structure of the precursor to the three major surface antigens of *Plasmodium falciparum* merozoites. *Nature* (*London*) 317, 270–273.

HOPP, T. P., AND WOODS, K. R. 1981. Prediction of protein antigenic determinants from amino acid sequences. *Proceedings of the National Academy of Science USA* 78, 3824–3828.

HUI, G. S. N., AND SIDDIQUI, W. A. 1987. Serum from Pf195 protected *Aotus* monkeys inhibit *Plasmodium falciparum* growth *in vitro*. *Experimental Parasitology* 64, 519–522.

KYTE, J., AND DOOLITTLE, R. 1982. A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology* 157, 105–132.

LYON, J. A., GELLER, R. A., HAYNES, J. D., CHULAY, J. D., AND WEBER, J. L. 1986. Epitope map and processing scheme for the 195,000 dalton surface glycoprotein of *Plasmodium falciparum* merozoites deduced from cloned overlapping segments of the gene. *Proceedings of the National Academy of Science USA* 83, 2989–2993.

MACKAY, M., GOMAN, M., BONE, N., HYDE, J. E., SCAIFE, J., CERTA, U., STUNNENBERG, H., AND BUJARD, H. 1985. Polymorphism of the precursor for the major surface antigens of *Plasmodium falciparum* merozoites studies at the genetic level. *EMBO Journal* 4, 3823–3829.

MANIATIS, T., FRITSCH, E. F., AND SAMBROOK, J. 1982. "Molecular Cloning. A Laboratory Manual," pp. 280–281. Cold Spring Harbor Laboratory, New York.

NUSSENZWEIG, V., AND NUSSENZWEIG, R. S. 1985. Circumsporozoite proteins of malaria parasites. *Cell* 42, 401–403.

PATARROYO, M. A., ROMERO, P., TORRES, M. L., CLAVIJO, P., MORENO, A., MARTINEZ, A., RODRIGUEZ, R., GUZMAN, F., AND CABEZAS, E. 1987. Induction of protective immunity against experimental infection with malaria using synthetic peptides. *Nature* (*London*) 328, 629–632.

PAULETTI, D., SIMMONDS, R., DREESMAN, G. R., AND KENNEDY, R. C. 1985. Application of a modified computer algorithm in determining potential antigenic determinants associated with the AIDS virus glycoprotein. *Analytical Biochemistry* 151, 541–546.

PERRIN, L. H., MERKLI, B., LOCHE, M., CHIZZOLINI, C., SMART, J., AND RICHLE, R. 1984. Antimalarial immunity in Saimiri monkeys. Immunization with surface components of asexual blood stages. *Journal of Experimental Medicine* 160, 441–451.

PETERSON, M. G., COPPEL, R. L., MCINTYRE, P., LANGFORD, C. J., WOODROW, G., BROWN, G. V., ANDERS, R. F., AND KEMP, D. J. 1988. Variation in the precursor to the major merozoite surface antigens of *Plasmodium falciparum*. *Molecular and Biochemical Parasitology* 27, 291–302.

SANGER, F., NICKLEN, S., AND COULSON, A. R. 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Science USA* 74, 5463–5467.

SIDDIQUI, W. A., AND PALMER, K. L. 1981. Propagation of malaria parasites *in vitro*. *In* "Advances in Cell Culture" (K. Maramorosch, Ed.), Vol. 1, pp. 182–212. Academic Press, New York.

SIDDIQUI, W. A., SCHNELL, J. V., AND GEIMAN, Q. M. 1972. A model *in vitro* system to test the susceptibility of human malarial parasites to antimalarial drugs. *American Journal of Tropical Medicine and Hygiene* 21, 392–399.

SIDDIQUI, W
G. S. N., (
S. P., CHA
ite surface
tects *Aotu*
parum mal
emy of Sci
SIDDIQUI, V
KRAMER, I
tial prote
vaccinated
heterologo
Medicine a
SIMPSON, L.,
SIMPSON,
AND STUA
cle (mitoch
lae and *Tr*
otide sequ
262, 6182–

SIDDIQUI, W. A., TAM, L. Q., KRAMER, K. J., HUI, G. S. N., CASE, S. E., YAMAGA, K. M., CHANG, S. P., CHAN, E. B., AND KAN, S.-C. 1987. Merozoite surface coat precursor protein completely protects *Aotus* monkeys against *Plasmodium falciparum* malaria. *Proceedings of the National Academy of Science USA* 84, 3014–3018.

SIDDIQUI, W. A., TAYLOR, D. W., KAN, S.-C., KRAMER, K., AND RICHMOND-CRUM. S. 1978. Partial protection of *Plasmodium falciparum*-vaccinated *Aotus trivirgatus* against a challenge of a heterologous strain. *American Journal of Tropical Medicine and Hygiene* 27, 1277–1278.

SIMPSON, L., NECKELMANN, N., DE LA CRUZ, B. F., SIMPSON, A. M., FEAGIN, J. E., JASMER, D. P., AND STUART, K. 1987. Comparison of the maxicircle (mitochondrial) genomes of *Leishmania tarentolae* and *Trypanosoma brucei* at the level of nucleotide sequence. *Journal of Biological Chemistry* 262, 6182–6196.

SNIDER, M. D. 1984. Biosynthesis of glycoproteins: Formation of N-linked oligosaccharides. *In* "Biology of Carbohydrates" (V. Ginsburg and P. Robbins, Eds.), Vol. 2, pp. 163–198. Wiley, New York.

TANABE, K., MACKAY, M., GOMAN, M., AND SCAIFE, J. G. 1987. Allelic dimorphism in a surface antigen gene of the malaria parasite *Plasmodium falciparum*. *Journal of Molecular Biology* 195, 273–287.

TRAGER, W., AND JENSEN, J. B. 1976. Human malaria parasites in continuous culture. *Science* 193, 673–675.

ULLRICH, A., BERMAN, C. H., DULL, T. J., GRAY, A., AND LEE, J. M. 1984. Isolation of the human insulin-like growth factor gene using a single synthetic DNA probe. *EMBO Journal* 3, 361–364.

WEBER, J. L., LEININGER, W. M., AND LYON, J. A. 1986. Variation in the gene encoding a major merozoite surface antigen of the human malaria parasite *Plasmodium falciparum*. *Nucleic Acids Research* 14, 3311–3323.

Molecular and Biochemical Parasitology, 36 (1989) 271–282
Elsevier

MOLBIO 01210

, Krakow. J.L., Eng-
7) A glycan-phospha-
D in human serum.

88) A phospholipase
ol anchor of cell-sur-
roc. Natl. Acad. Sci.

irt, G.W. and Eng-
or glycan assembly:
dylinositol anchor of
coprotein. Cell 56.

'.T. and Hart, G.W.
ace glycoprotein of
63, 17697–17705.

Cloning and analysis of the gene encoding the 230-kilodalton merozoite surface antigen of *Plasmodium yoelii*

Alan P. Lewis

Department of Molecular Biology, Wellcome Biotech, Beckenham, Kent, U.K.

(Received 31 March 1989; accepted 16 May 1989)

The complete nucleotide sequence of the gene for the 230-kDa precursor to the major merozoite surface antigens (PMMSA) of *Plasmodium yoelii* YM has been determined. A single open reading frame of 5316 bp encodes a protein of calculated molecular mass 197 000. The deduced amino acid sequence contains potential signal peptide and membrane anchor sequences of 19 and 18 residues, respectively, and a region of six tandemly repeated tetrapeptides, Gly-Ala-Val-Pro. Comparison with the 195-kDa *Plasmodium falciparum* analogue (Pf195) at the amino acid level indicated an overall homology of approximately 30%, with areas of as high as 60% conservation. The tripeptide repeats present near the N-terminus of the Pf195 antigen are absent from the Py230 sequence. The PMMSA can be divided into 22 blocks based upon interspecies amino acid conservation.

Key words: *Plasmodium yoelii*; Malaria; Precursor to the major merozoite surface antigens; Nucleic acid sequence

Introduction

The precursor to the major merozoite surface antigens (PMMSA) has been proposed as a candidate for a vaccine directed against the asexual erythrocytic stage of malaria [1]. This polypeptide has been identified in human [2,3], simian [4], and rodent [5–7] malarial species, ranging in molecular mass from 185 to 250 kDa. The precursor is synthesised during intraerythrocytic development of the parasite, and is processed into a number of discrete fragments during merozoite maturation [2,8–11]. PMMSA fragments have been detected on the surface coat of the merozoite [12–15] and a possible role for the protein has been suggested in erythrocyte invasion [16].

The PMMSA has a size of 230 kDa in the murine malaria *Plasmodium yoelii* (Py230) [5]. Purified Py230, and a monoclonal antibody which recognises a C-terminal epitope on this protein, have both been shown to protect mice from challenge infection with *P. yoelii* [5,17,18]. A role for cell-mediated immunity has also been implicated in the response observed to the whole antigen [19]. The analogous protein in the human malaria *Plasmodium falciparum* has a size in the range 185–205 kDa (Pf195) [2]. The complete protein [20–22], or synthetic peptides derived from this antigen [23,24], have been used to produce partial or complete protection against challenge infection in non-human vaccine trials. Furthermore, a polymeric synthetic hybrid protein, based upon a mixture of three synthetic peptides including a derivative of Pf195, has been found to induce protective immunity in humans [25]. These results reinforce the potential of the PMMSA as a candidate for a vaccine against the malarial asexual blood stage, and emphasize the need to develop an experimental model system for this antigen in order to analyse in more detail the mechanisms involved in protective immunity.

Correspondence address: Alan P. Lewis, Dept. of Molecular Biology, Wellcome Biotech, Langley Court, Beckenham, Kent, BR3 3BS, U.K.

Note: Nucleotide sequence data reported in this paper have been submitted to the GenBank™ Data Bank with the accession number J04668.

Abbreviations: PMMSA, precursor to the major merozoite surface antigens; Py230, 230-kDa merozoite antigen of *P. yoelii*; Pf195, 195-kDa merozoite antigen of *P. falciparum*.

The 3′ portion of the gene for the *P. yoelii* 17XL 230-kDa protein was recently published [26]. This study describes the cloning and sequencing of the complete Py230 gene from the YM strain of *P. yoelii*, and a comparison of the sequence with the *P. falciparum* analogue, Pf195.

Materials and Methods

Preparation of parasite genomic DNA. CD1 mice were infected with the YM strain of *P. yoelii*, and lymphocytes removed on day 1 post-infection by subcutaneous injection with 150 μl of 40 mg ml^{-1} cyclophosphamide in 0.85% NaCl [27]. Parasitised blood was collected on day 3 post-infection, when parasitaemias averaged 50%, and parasite DNA was prepared as previously described [28.29].

Construction of genomic DNA libraries. Three libraries were constructed from the *P. yoelii* genomic DNA. All restriction endonuclease digests were under conditions recommended by the manufacturer. Library (1): parasite DNA was digested with mung bean nuclease (Pharmacia) in 35% formamide as described [30], and ligated to phosphorylated *Eco*RI linkers (Pharmacia). After digestion with *Eco*RI (Amersham), excess linkers were removed using a NACS PREPAC column (BRL). The DNA was ligated into the *Eco*RI site of λgt11 (Stratagene), and a genomic library constructed by in vitro packaging (Stratagene). Library (2): parasite DNA was digested with *Dra*I (NBL) and ligated into the *Sma*I site of pUC9 (Pharmacia) treated with calf intestinal alkaline phosphatase [28]. A genomic library was constructed by transformation of Max Efficiency DH5α competent cells (BRL). Library (3): parasite DNA was digested with *Eco*RI and a genomic library constructed, as above, using *Eco*RI-cut pUC9.

Screening of genomic libraries. Libraries were screened using synthetic oligonucleotides made on a Biosearch Sam One DNA synthesiser (New Brunswick). Library (1) was screened using probes A (a 26-mer, 5′-GAAGGTAATA-CATGTGTAGAAAATAA-3′ corresponding to nucleotides 1807–1832 in ref. 26) and B (a 26-mer, 5′-TTTCTTTAACAAGAGAAGAGAA-GCTG-3′ corresponding to nucleotides 285–310 in ref. 26); library (2) was screened using probe B, and library (3) using probe C (an 18-mer, 5′-AAACAAAGATGCTTTAAG-3′ corresponding to nucleotides 2685–2702 in Fig. 2). Bacteriophage plaques and bacterial colonies were lifted on to Hybond-N nylon filters (Amersham) following the manufacturer's instructions. and screened with ^{32}P-labelled oligonucleotides [31] as previously described [32]. Bacteriophage λ and pUC9 plasmid DNA was isolated from positive clones as described [28,33].

DNA sequencing. The insert of the λgt11 recombinant was subcloned into the *Eco*RI site of pUC9. Regions of the genomic DNA pUC9 clones were sequenced in both directions by plasmid priming following the dideoxy chain termination method [34], according to the Sequenase kit (USB) protocol. Sequences were analysed on the Wellcome Biotech computer system with the aid of the programs IALIGN (National Biomedical Research Foundation) and DIAGON [35].

Results

Isolation of P. yoelii YM Py230 clones. The 3′ portion of the gene for the Py230 antigen from *P. yoelii* 17XL was published recently [26]. Two oligonucleotide probes, A and B, were synthesised. corresponding to nucleotides 1807–1832 and 285–310, respectively, from the published sequence. These regions were chosen as they possessed high nucleic acid homology to the corresponding sections of the Wellcome Pf195 sequence [36]. Approximately 7×10^4 phage from library (1) were screened with probe A, and 10 positive clones were detected. After additional rounds of screening with probes A and B, one recombinant, λPyM4.3, remained positive. λPyM4.3 contained an *Eco*RI insert of 4.3 kb (Fig. 1).

Probe B was used to screen approximately 4×10^4 recombinants of library (2). One positive clone was isolated, pPyD1.7, which was found to possess a *Dra*I insert of 1.7 kb (Fig. 1). From sequence analysis, oligonucleotide probe C was synthesised and used to screen library (3). Approximately 2×10^4 recombinants were probed.

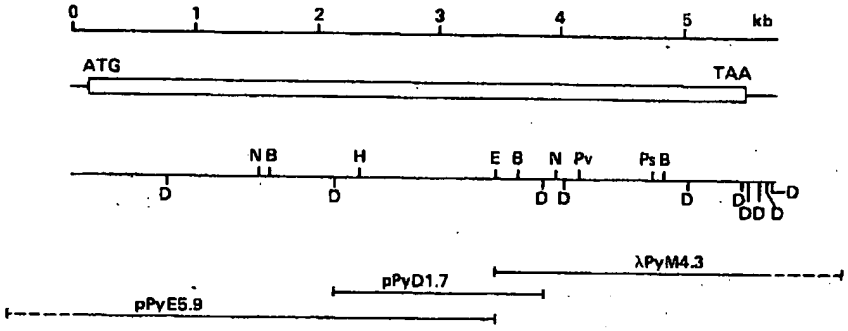


Fig. 1. Physical map of the Py230 gene from the YM strain of *P. yoelii*. The upper map indicates the positions of the start and stop codons, and the length of the coding region. The middle section represents a partial restriction map of the gene. Restriction enzyme sites indicated are: B, *Bgl*II; D, *Dra*I; E, *Eco*RI; H, *Hin*cII; N, *Nde*I; Ps, *Pst*I and Pv, *Pvu*II. The lower map shows the genomic DNA clones used for sequence analysis.

and 20 positives isolated. A genomic DNA clone, pPyE5.9, that was purified, contained a 5.9-kb *Eco*RI insert (Fig. 1).

A fragment from λPyM4.3 was used to probe an RNA blot of total *P. yoelii* YM RNA. A 7.3-kb RNA species was detected (data not shown); the size expected for a transcript encoding a 230-kDa protein (including 5′ leader and 3′ non-coding sequences).

Nucleotide sequence of the P. yoelii YM Py230 gene. The λPyM4.3 insert was subcloned into *Eco*RI-cut pUC9. The DNA sequence was determined for a region overlapping the inserts from the three recombinant clones, spanning 5775 nucleotides (Fig. 2). A methionine start codon at nucleotide 190 is followed by a single open reading frame of 5316 bp terminating with the first stop codon at nucleotide 5505. The A+T content is high, with an average of 69% within the coding region and 85% for the 5′ and 3′ untranslated sequences. This is consistent with levels found for the entire *P. yoelii* genome [37]. The open reading frame encodes a polypeptide of 1772 amino acid residues with a calculated size of 197 kDa, thus smaller than the 230 kDa determined using sodium dodecyl sulphate-polyacrylamide gel electrophoresis. A number of other malarial antigens [36,38–41], however, also exhibit such discrepancies, a property thought to be related to the repetitive regions of these polypeptides. The Py230 amino acid sequence contains six tandem repeats of the tetrapeptide Gly-Ala-Val-Pro, which may account for a similar discrepancy in this antigen. At the N-terminus of the polypeptide is a putative signal peptide of 19 amino acids, and at the C-terminus a potential 18-amino-acid hydrophobic membrane anchor. The sequence contains 20 cysteine residues, of which 10 are situated within the C-terminal 110 amino acids. Of the remaining 10 cysteines, 8 appear to be positioned as 4 pairs, based upon their linear proximities. There are also 11 potential *N*-glycosylation sites (Asn-X-Ser/Thr, where X can be any amino acid with the probable exclusion of proline [42]) scattered throughout the molecule. The sequence of the C-terminal 2310 nucleotides is identical to that published for the Py230 gene from *P. yoelii* 17XL [26]. The virulent YM and 17XL strains were originally derived from a common ancestor, the uncloned avirulent isolate, 17X [43,44].

Comparisons between the Py230 and Pf195 sequences. The *P. yoelii* YM Py230 amino acid sequence was aligned with the Wellcome strain Pf195 sequence [36] by computer analysis (Fig. 3; a revised Wellcome Pf195 sequence was used which has been submitted to the GenBank and EMBL databases). An overall homology of 31% was determined, with particular regions exhibiting as much as 60% conservation. 14 of the 20 cysteines within the Py230 sequence are located at positions similar to those in Pf195, including all 10 cysteines at the C-terminus. None of the *N*-glycosylation sites present in either polypeptide, however, are conserved. This may reflect the ob-

274

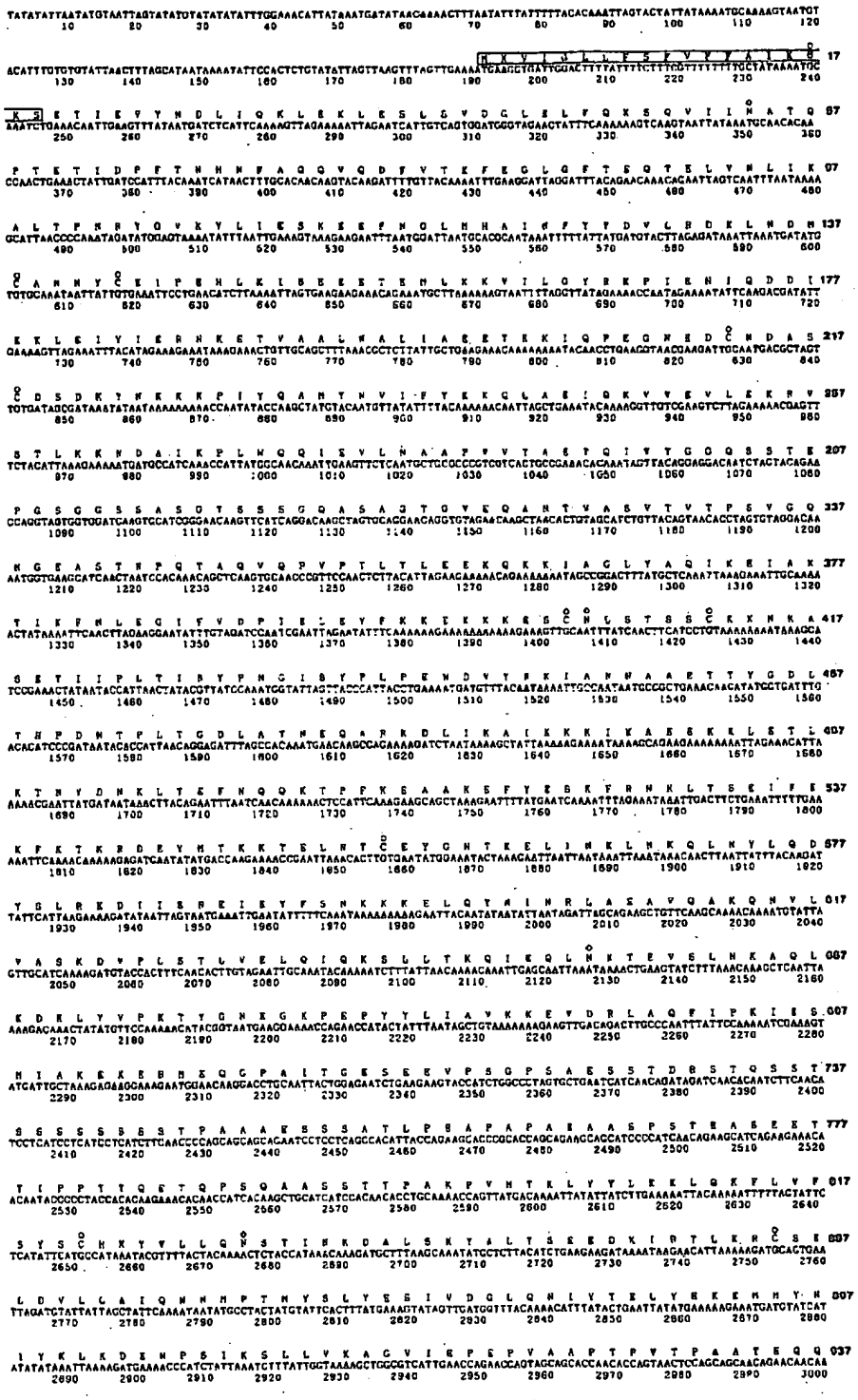


Fig. 2. Nucl
peptide and
Positi

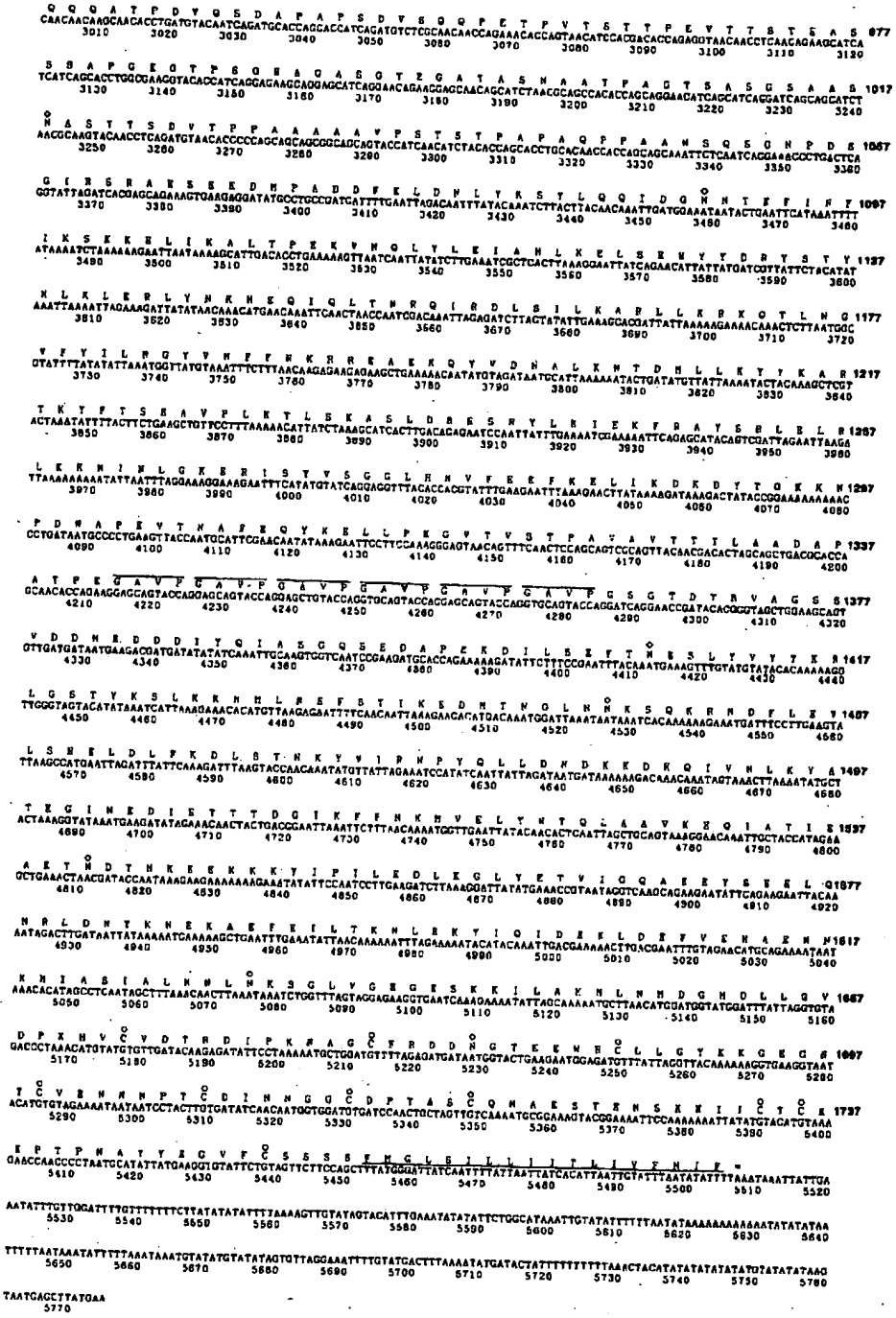


Fig. 2. Nucleotide sequence of the Py230 gene, and the deduced amino acid sequence of the open reading frame. Putative signal peptide and membrane anchor sequences are boxed and underlined, respectively, whilst the tetrapeptide repeats are overlined. Positions of cysteine residues are denoted by open circles, and potential *N*-glycosylation sites by open diamonds.

276

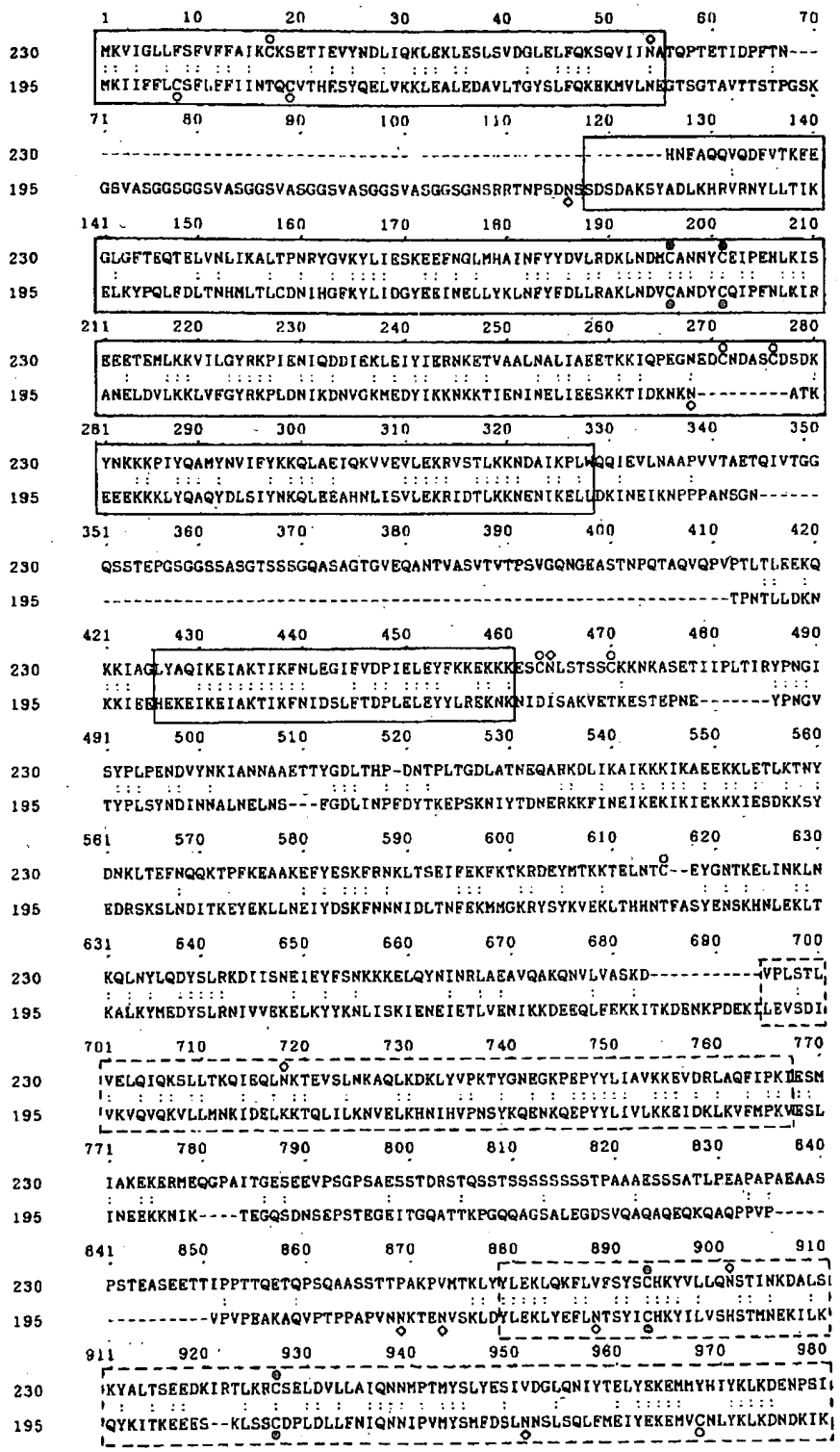


Fig. 3. Com
using the [AI
tween sequen
sylation sites
amino acids b

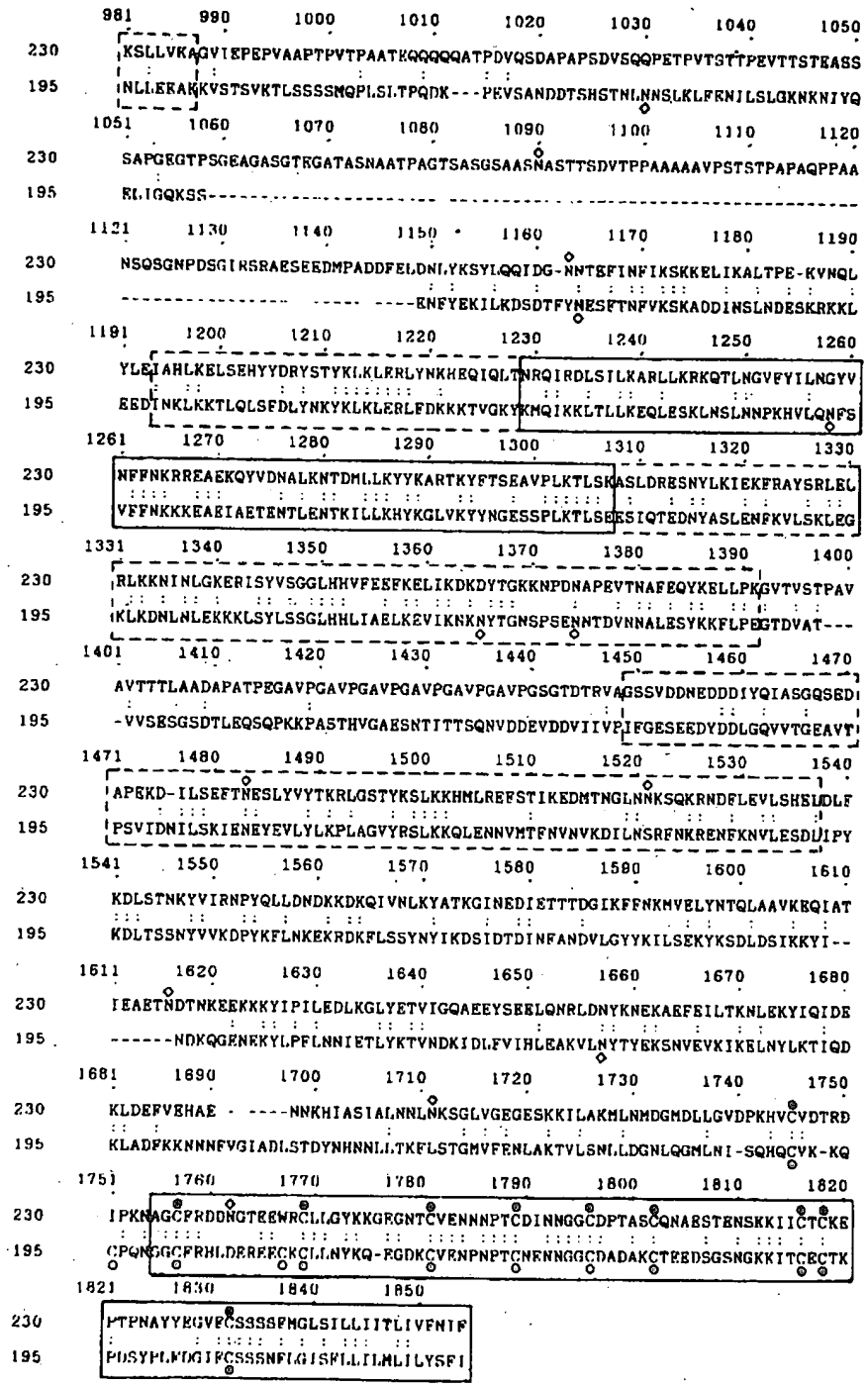
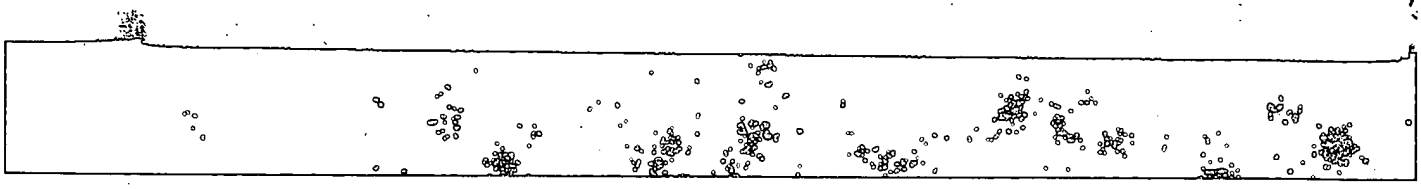
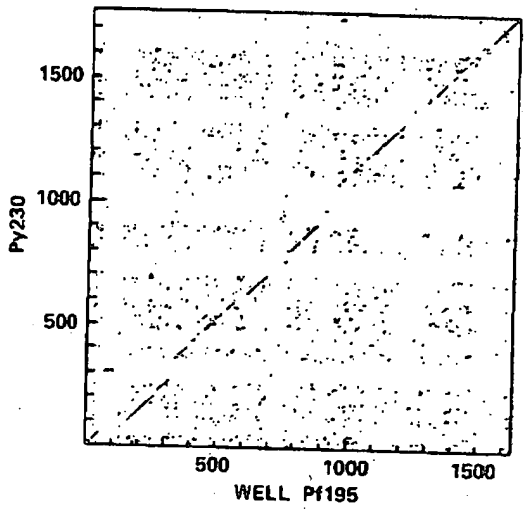


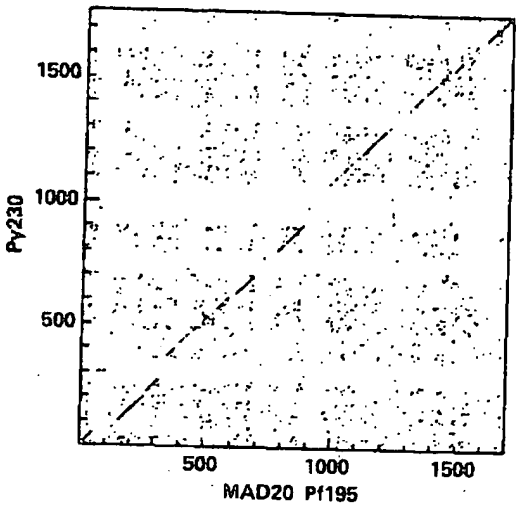
Fig. 3. Comparison of the amino acid sequence of Py230 with Pf195 from the Wellcome strain [36]. Alignment was carried out using the IALIGN program (National Biomedical Research Foundation). Positions of cysteine residues that are conserved between sequences are indicated by filled circles, and those that are not conserved by open circles. Positions of potential *N*-glycosylation sites are denoted by open diamonds. The positions of previously determined Pf195 blocks based upon conservation of amino acids between different Pf195 alleles [47] are shown. Conserved blocks are boxed by unbroken lines, semi-conserved blocks by broken lines, and variable blocks remain unboxed.



servation that the Py230 antigen is lacking in glycosylation [8], and that the glycosylation identified in the Pf195 may be confined to a glycolipid anchor identified at the C-terminus [45,46]. The Py230 sequence, when compared to Pf195, possesses two large 'inserted' blocks of amino acids in the central region and near to the N-terminus of the polypeptide. There is also a large deletion of residues following the putative signal peptide, which spans the tripeptide repeats (Ser-X-X) present in the Wellcome Pf195.

The Py230 amino acid sequence was compared to the two Pf195 allelic variants from the Wellcome and MAD20 [47] strains of *P. falciparum*, using the DIAGON computer program of Staden [35] (Fig. 4). Regions of conservation between Py230 and the Wellcome Pf195 allele were found to be similarly conserved when the MAD20 Pf195 allele was compared. The Py230 and Wellcome Pf195 sequences have also been compared to the published sequence for a portion of the 200-kDa PMMSA from the human malaria *Plasmodium vivax* [3] (data not shown). This aligns with the region 117–794 amino acid residues of Py230. It





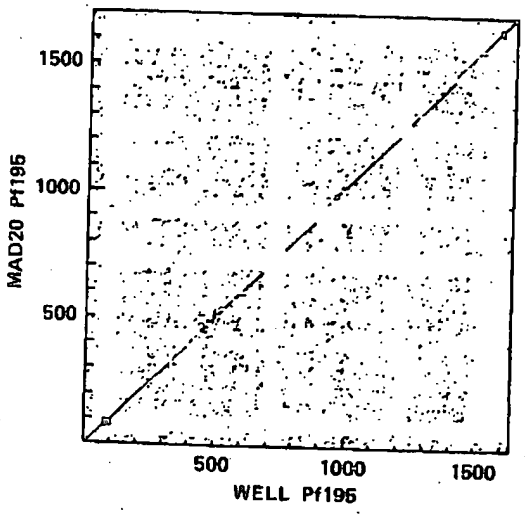


Fig. 4. Comparisons of the Py230, Wellcome Pf195 [36] and MAD20 Pf195 [47] amino acid sequences. Analysis was by the DIAGON program of Staden [35], using a proportional algorithm (proportional score 132, span length 11). The axes of the plots represent the appropriate amino acid sequences numbered from their N- to C-termini.

was found
tween the
closely mi
of the oth

Discussion

The dat
and sequei
kDa mero
protein ha
the 195-kD
ilar putati
chor seque
ologies thr
The level o
Py230 as a
PMMSA.

Certain
tween the
P. yoelii,
parasites ar
ary groups,
position an
laria genes
situates in a
amino acids
the three e
gues that the
these region

The *P. fal*
viously been
conservation
Pf195 alleles
ified as eithe
ogy), semi-cc
or variable (
the alignmen
may not nec
tion of amin

Fig. 5. Represen
The sequence is
possessing be

was found that conservation of amino acids between the PMMSAs of any two of the species closely mirrored the homologies observed for each of the other species comparisons.

Discussion

The data presented here describe the cloning and sequencing of the complete gene for the 230-kDa merozoite antigen from *P. yoelii* YM. The protein has a structure closely resembling that of the 195-kDa PMMSA of *P. falciparum*, with similar putative signal peptide and membrane anchor sequences, and extensive amino acid homologies throughout the length of the molecule. The level of similarity consolidates the use of the Py230 as an experimental model system for the PMMSA.

Certain protein sequences are conserved between the PMMSAs of the three malarial species *P. yoelii*, *P. falciparum* and *P. vivax*. Malaria parasites are thought to fall into three evolutionary groups, based upon genomic DNA base composition and sequence similarities between malaria genes [48,49], and each of the above species situates in a separate group. The conservation of amino acids observed between the PMMSAs of the three evolutionarily distant species thus argues that there are certain constraints placed upon these regions of the polypeptide.

The *P. falciparum* PMMSA sequence has previously been divided into 17 blocks based upon conservation of amino acids between different Pf195 alleles [47]. These regions have been classified as either conserved (more than 87% homology), semi-conserved (areas of patchy homology) or variable (extensive divergence). As shown by the alignment in Fig. 3, however, these divisions may not necessarily reflect the actual conservation of amino acid sequences between malarial species. Certain 'variable' blocks from the Pf195 allelic analysis, such as amino acids 385–608 from the Wellcome sequence, can be seen to contain regions of close homology when Py230/Pf195 comparisons are made. By contrast, 'conserved' blocks can possess areas of comparatively little homology. DIAGON analysis indicates that any homologies observed between Py230 and Pf195 can also be reproduced in comparisons between the different Pf195 alleles from the Wellcome and MAD20 strains, even within the 'variable' blocks (Fig. 4). This result suggests that the interspecies conservation can be taken as a good indication of the PMMSA regions that are associated with essential structural and/or functional roles.

The Py230 amino acid sequence can be divided into 22 different blocks based upon interspecies conservation (Fig. 5). The blocks are classified thus; (a) conserved (possessing greater than 45% homology); (b) semi-conserved (between 20 and 45% homology); and (c) variable (less than 20% homology and frequently containing large deletions or insertions of amino acids). All of the conserved cysteine residues are found within conserved blocks, thus suggesting important structural functions for these amino acids. The semi-conserved blocks often contain sequences of low homology interspersed with small regions of high conservation, and the latter sequences again probably signify amino acids of physical importance to the protein. The conserved and semi-conserved blocks are essentially α-helical in structure with the variable regions consisting of randomly coiled hydrophilic amino acids (data not shown). Such data suggest that the conserved blocks represent sequences internal to the protein, with the variable regions positioned on the extremities at the apices of adjacent α-helices. This would allow for the sequence variability observed. Exceptions to the rule are the repeat re-

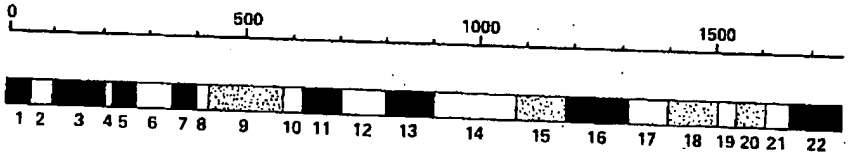


Fig. 5. Representation of the Py230 protein sequence based upon amino acid conservation between the Py230 and Pf195 antigens. The sequence is divided into conserved blocks of greater than 45% amino acid homology (black boxes), semi-conserved blocks possessing between 20 and 45% homology (shaded boxes), and variable blocks of less than 20% homology (open boxes).

gions, Gly-Ala-Val-Pro and Ser-X-X, present in variable blocks of Py230 and Pf195, respectively. These are hydrophobic in each case, suggesting that they may be positioned on the interior of the protein. It has been proposed that tandem repeats in malaria antigens are likely to possess some major role, perhaps playing a critical part in the specific functioning of the antigen, or acting as an immunological decoy directing immune responses away from more functionally important regions of the polypeptide [49,50]. The assumed deep-seated positions of these structures, however, combined with the fact that repeats present in the PMMSA of one species are lacking in that of the other species, argues against both of these suggested roles. Indeed, a Pf195 antigen has been described that is totally lacking in repeats [51]. Whether the tandem repeats of the PMMSA possess an important function thus remains unclear.

The external positioning of the variable regions within the PMMSA makes them potentially highly immunogenic. The very variability of such sequences, however, argues against them as being effective in a malarial vaccine. Two T-cell epitopes have been identified within the N-terminal portion of the Pf195 [52], and both are situated in a conserved region of the PMMSA (Fig. 5; block 3). The conserved blocks could perhaps be used as a basis to limit the search for valuable PMMSA B- and T-cell epitopes, as such epitopes, within these areas, may be expected to show little or no intraspecies antigenic variation.

Acknowledgements

I would like to thank Steve Nicholls for assistance with computing, Hugh Spence for oligonucleotide synthesis, Dr. Mike Lockyer for critical reading of the manuscript, Dr. Tony Holder for helpful discussions, and Pam Sparks for secretarial help. This work was supported by a United Nations Development Program/World Bank/World Health Organization Special Programme for Research and Training in Tropical Diseases.

References

1 Holder, A.A. (1988) The precursor to major merozoite surface antigens: structure and role in immunity. Prog. Allergy. 41, 72–97.
2 Holder, A.A. and Freeman, R.R. (1982) Biosynthesis and processing of a *Plasmodium falciparum* schizont antigen recognized by immune serum and a monoclonal antibody. J. Exp. Med. 156, 1528–1538.
3 Del Portillo, H.A., Gysin, J., Mattei, D.M., Khouri, E., Udagama, P.V., Mendis, K.N. and David, P.H. (1988) *Plasmodium vivax*: cloning and expression of a major blood-stage surface antigen. Exp. Parasitol. 67, 346–353.
4 Epstein, N., Miller, L.H., Kaushel, D.C., Udeinya, I.J., Rener, J., Howard, R.J., Asofsky, R., Aikawa, M. and Hess, R.L. (1981) Monoclonal antibodies against a specific surface determinant on malarial (*Plasmodium knowlesi*) merozoites block erythrocytic invasion. J. Immunol. 127, 212–217.
5 Holder, A.A. and Freeman, R.R. (1981) Immunization against blood-stage rodent malaria using purified parasite antigens. Nature 294, 361–364.
6 Boyle, D.B., Newbold, C.I., Smith, C.C. and Brown, K.N. (1982) Monoclonal antibodies that protect in vivo against *Plasmodium chabaudi* recognize a 250000-dalton parasite polypeptide. Infect. Immun. 38, 94–102.
7 Newbold, C.I., Boyle, D.B., Smith, C.C. and Brown, K.N. (1982) Identification of a schizont and species specific surface glycoprotein on erythrocytes infected with rodent malarias. Mol. Biochem. Parasitol. 5, 45–54.
8 Holder, A.A. and Freeman, R.R. (1984) Characterisation of a high molecular weight protective antigen of *Plasmodium yoelii*. Parasitology 88, 211–219.
9 David, P.H., Hadley, T.J., Aikawa, M. and Miller, L.H. (1984) Processing of a major parasite surface glycoprotein during the ultimate stages of differentiation in *Plasmodium knowlesi*. Mol. Biochem. Parasitol. 11, 267–282.
10 Hall, R., Osland, A., Hyde, J.E., Simmons, D.L., Hope, I.A. and Scaife, J. (1984) Processing, polymorphism and biological significance of P190, a major surface antigen on the erythrocytic forms of *Plasmodium falciparum*. Mol. Biochem. Parasitol. 11, 61–80.
11 Lyon, J.A., Geller, R.H., Haynes, J.D., Chulay, J.D. and Weber, J.L. (1986) Epitope map and processing scheme for the 195000-dalton surface glycoprotein of *Plasmodium falciparum* merozoites deduced from cloned overlapping segments of the gene. Proc. Natl. Acad. Sci. USA 83 2989–2993.
12 Atkinson, C.T., Aikawa, M., Fujino, T., Tam, L.Q., Hui, G.S.N. and Siddiqui, W.A. (1987) Ultrastructural localization of protective and non-protective *Plasmodium falciparum* proteins using serum samples from vaccinated *Aotus* monkeys. J. Parasitol. 73, 1235–1240.
13 Howard, R.J., Lyon, J.A., Diggs, C.L., Haynes, J.D., Leech, J.H., Barnwell, J.W., Aley, S.B., Aikawa, M. and Miller, L.H. (1984) Localization of the major *Plasmodium falciparum* glycoprotein on the surface of mature intraer-

ythrocytic trophozoites and schizonts. Mol. Biochem. Parasitol. 11, 349–362.
14 Freeman, R.R. and Holder, A.A. (1983) Surface antigens of malaria merozoites. A high molecular weight precursor is processed to an 83 000-mol.wt form expressed on the surface of *Plasmodium falciparum* merozoites. J. Exp. Med. 158, 1647–1653.
15 Heidrich, H.-G., Matzner, M., Miettinen-Baumann, A. and Strych, W. (1986) Immunoelectron microscopy shows that the 80 000-dalton antigen of *Plasmodium falciparum* merozoites is localised in the surface coat. Z. Parasitenkd. 72, 681–683.
16 Hadley, T.J., Klotz, F.W. and Miller, L.H. (1986) Invasion of erythrocytes by malaria parasites: a cellular and molecular overview. Annu. Rev. Microbiol. 40, 451–477.
17 Freeman, R.R. and Holder, A.A. (1983) Characteristics of the protective response of BALB/c mice immunized with a purified *Plasmodium yoelii* schizont antigen. Clin. Exp. Immunol. 54, 609–616.
18 Majarian, W.R., Daly, T.M., Weidanz, W.P. and Long, C.A. (1984) Passive immunization against murine malaria with an IgG3 monoclonal antibody. J. Immunol. 132, 3131–3137.
19 Playfair, J.H.L., De Souza, J.B., Freeman, R.R. and Holder, A.A. (1985) Vaccination with a purified blood-stage malaria antigen in mice: correlation of protection with T cell mediated immunity. Clin. Exp. Immunol. 62, 19–23.
20 Perrin, L.H., Merkli, B., Loche, M., Chizzolini, C., Smart, J. and Richle, R. (1984) Antimalarial immunity in *Saimiri* monkeys. Immunization with surface components of asexual blood stages. J. Exp. Med. 160, 441–451.
21 Hall, R., Hyde, J.E., Goman, M., Simmons, D.L., Hope, I.A., Mackay, M., Scaife, J.,Merkli, B., Richle, R. and Stocker, J. (1984) Major surface antigen gene of a human malaria parasite cloned and expressed in bacteria. Nature 311, 379–382.
22 Siddiqui, W.A., Tam, L.Q., Kramer, K.J., Hui, G.S.N., Case, S.E., Yamaga, K.M., Chang, S.P., Chan, E.B.T. and Kan, S.-C. (1987) Merozoite surface coat precursor protein completely protects *Aotus* monkeys against *Plasmodium falciparum* malaria. Proc. Natl. Acad. Sci. USA 84, 3014–3018.
23 Cheung, A., Leban, J., Shaw, R.A., Merkli, B., Stocker, J., Chizzolini, C., Sander, C. and Perrin, L.H. (1986) Immunization with synthetic peptides of a *Plasmodium falciparum* surface antigen induces antimerozoite antibodies. Proc. Natl. Acad. Sci. USA 83, 8328–8332.
24 Patarroyo, M.E., Romero, P., Torres, M.L., Clavijo, P., Moreno, A., Martinez, A., Rodriguez, R., Guzman, F. and Cabezas, E. (1987) Induction of protective immunity against experimental infection with malaria using synthetic peptides. Nature 328, 629–632.
25 Patarroyo, M.E., Amador, R., Clavijo, P., Moreno, A., Guzman, F., Romero, P., Tascon, R., Franco, A., Murillo, L.A., Ponton, G. and Trujillo, G. (1988) A synthetic vaccine protects humans against challenge with asexual blood stages of *Plasmodium falciparum* malaria. Nature 332, 158–161.
26 Burns Jr., J.M., Daly, T.M., Vaidya, A.B. and Long, C.A. (1988) The 3′ portion of the gene for a *Plasmodium yoelii* merozoite surface antigen encodes the epitope recognized by a protective monoclonal antibody. Proc. Natl. Acad. Sci. USA 85, 602–606.
27 Berenbaum, M.C. (1975) The clinical pharmacology of immunosuppressive agents. In: Clinical Aspects of Immunology, 3rd ed., (Gell, P.G.H., Coombs, R.R.A. and Lachmann, P.J., eds.), pp. 689–758, Blackwell, Oxford, London, Edinburgh and Melbourne.
28 Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
29 Martin, W.J., Finerty, J. and Rosenthal, A. (1971) Isolation of *Plasmodium berghei* (Malaria) parasites by ammonium chloride lysis of infected erythrocytes. Nature New Biol. 233, 260–261.
30 McCutchan, T.F., Hansen, J.L., Dame, J.B. and Mullins, J.A. (1984) Mung bean nuclease cleaves *Plasmodium* genomic DNA at sites before and after genes. Science 225, 625–628.
31 Woods, D. (1984) Oligonucleotide screening of cDNA libraries. Focus 6 (3), 1–2.
32 Church, G.M. and Gilbert, W. (1982) Genomic sequencing. Proc. Natl. Acad. Sci. USA 81, 1991–1995.
33 Pratt, J.M. (1984) Coupled transcription-translation in prokaryotic cell-free systems. In: Transcription and translation: a practical approach (Hames, B.D. and Higgins, S.J., eds.), pp. 179–209, IRL Press, Oxford and Washington, DC.
34 Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74, 5463–5467.
35 Staden, R. (1982) An interactive graphics program for comparing and aligning nucleic acid and amino acid sequences. Nucleic Acids Res. 10, 2951–2961.
36 Holder, A.A., Lockyer, M.J., Odink, K.G., Sandhu, J.S., Riveros-Moreno, V., Nicholls, S.C., Hillman, Y., Davey, L.S., Tizard, M.L.V., Schwarz, R.T. and Freeman, R.R. (1985) Primary structure of the precursor to the three major surface antigens of *Plasmodium falciparum* merozoites. Nature 317, 270–273.
37 Chance, M.L., Warhurst, D.C., Baggaley, V.C. and Peters, W. (1972) Preparation and characterisation of DNA from rodent malarias. Trans. R. Soc. Trop. Med. Hyg. 66, 3–4.
38 Ozaki, L.S., Svec, P., Nussenzweig, R.S., Nussenzweig, V. and Godson, G.N. (1983) Structure of the *Plasmodium knowlesi* gene coding for the circumsporozoite protein. Cell 34, 815–822.
39 Eichinger, D.J., Arnot, D.E., Tam, J.P., Nussenzweig, V. and Enea, V. (1986) Circumsporozoite protein of *Plasmodium berghei*: gene cloning and identification of the immunodominant epitopes. Mol. Cell. Biol. 6, 3965–3972.
40 Stahl, H.D., Kemp, D.J., Crewther, P.E., Scanlon, D.B., Woodrow, G., Brown, G.V., Bianco, A.E., Anders, R.F. and Coppel, R.L. (1985) Sequence of a cDNA encoding a small polymorphic histidine- and alanine-rich protein from *Plasmodium falciparum*. Nucleic Acids Res. 13, 7837–7846.
41 Favaloro, J.M., Coppel, R.L., Corcoran, L.M., Foote,

S.J., Brown, S.V., Anders, R.F. and Kemp, D.J. (1986) Structure of the RESA gene of *Plasmodium falciparum*. Nucleic Acids Res. 14, 8265–8277.
42 Von Heijne, G. (1987) Sequence Analysis in Molecular Biology: Treasure Trove or Trivial Pursuit. Academic Press, London.
43 Yoeli, M., Hargreaves, B.J., Carter, R. and Walliker, D. (1975) Sudden increase in virulence in a strain of *Plasmodium berghei yoelii*. Ann. Trop. Med. Parasitol. 69, 173–178.
44 Weinbaum, F.I., Evans, C.B. and Tigelaar, R.E. (1976) An in vitro assay for T cell immunity to malaria in mice. J. Immunol. 116, 1280–1283.
45 Haldar, K., Ferguson, M.A.J. and Cross, G.A.M. (1985) Acylation of a *Plasmodium falciparum* merozoite surface antigen via *sn*-1,2-diacyl glycerol. J. Biol. Chem. 260, 4969–4974.
46 Schwarz, R.T., Lockyer, M.J. and Holder, A.A. (1987) Aspects of the post-translational modification of a major *Plasmodium falciparum* merozoite surface antigen. In: Host-Parasite Cellular and Molecular Interactions in Protozoal Infections (Chang, K.P. and Snary, D., eds.), pp. 275–280, Springer-Verlag, Heidelberg.
47 Tanabe, K., Mackay, M., Goman, M. and Scaife, J.G. (1987) Allelic dimorphism in a surface antigen gene of the malaria parasite *Plasmodium falciparum*. J. Mol. Biol. 195, 273–287.
48 McCutchan, T.F., Dame, J.B., Miller, L.H. and Barnwell, J. (1984) Evolutionary relatedness of *Plasmodium* species as determined by the structure of DNA. Science 225, 808–811.
49 Weber, J.L. (1988) Molecular biology of malaria parasites. Exp. Parasitol. 66, 143–170.
50 Kemp, D.J., Coppel, R.L. and Anders, R.F. (1987) Repetitive proteins and genes of malaria. Annu. Rev. Microbiol. 41, 181–208.
51 Certa, U., Rotmann, D., Matile, H. and Reber-Liske, R. (1987) A naturally occurring gene encoding the major surface antigen precursor p190 of *Plasmodium falciparum* lacks tripeptide repeats. EMBO J. 6, 4137–4142.
52 Crisanti, A., Muller, H.-M., Hilbich, C., Sinigaglia, F., Matile, H., McKay, M., Scaife, J., Beyreuther, K. and Bujard, H. (1988) Epitopes recognized by human T cells map within the conserved part of the GP190 of *P. falciparum*. Science 240, 1324–1326.

Molecular
Elsevier

MOLBIO

Precu

Key words:

The pr
antigen (
candidate
Plasmodi
ported th
gene enc
rodent m
the gene
yoelii PM
tope recc
tibody [2
been loca
antigen (
portant e
we have
reported
P. yoel
size-selec
subgenon
*Nde*I frag
insert. Py
was isolat

Correspond
and Immun
Sts., Philade

Note: Nucl
been submit
sion numbe

Abbreviatio
surface antig

0166-6851/8

ol. 3, 257–265.
ti, D.V. (1985) Purifi-
bifunctional thymidy-
ase from methotrexate-
emistry 24, 678–686.
K. (1984) Purification
olyzing enzyme from
hem. 48, 505–511.
owsky, A., and Bever-
g-resistant *Leishmania*
-10-methyl-4-deoxy-4-
of methotrexate polyg-
0–15966.
Molnar, D.A. (1961)
ubstrate and cofactor
s by cell-free extracts
36, 2534–2543.
ann, M.P. (1964) The
ipounds from hydrox-
ate. J. Biol. Chem.

Molecular and Biochemical Parasitology, 43 (1990) 231–244
Elsevier

MOLBIO 01416

Molecular cloning and sequence analysis of the gene encoding the major merozoite surface antigen of *Plasmodium chabaudi chabaudi* IP-PC1

Willy Deleersnijder, Diana Hendrix, Najib Bendahman, Josée Hanegreefs, Lea Brijs, Cécile Hamers-Casterman and Raymond Hamers

Laboratorium Algemene Biologie, Instituut voor Moleculaire Biologie, Vrije Universiteit Brussel, Brussels, Belgium

(Received 12 March 1990; accepted 27 June 1990)

The complete nucleotide sequence of the gene encoding the precursor to the major merozoite surface antigens of *Plasmodium chabaudi chabaudi* strain IP-PC1 has been determined. A single open reading frame was detected, that coded for a protein of 199 kDa. The encoded protein (p199) contains putative signal and membrane anchor sequences and shows a clustering of Cys residues in the last 120 amino acids. Incompletely conserved tandem repeat oligopeptides are present at different positions in the molecule. P199 shows 69% overall homology to the analogous antigen in *Plasmodium yoelii yoelii* strain YM. The divergence between these antigens is largely confined to 4 areas where a number of insertions and/or deletions have occurred. All repeats occur in these divergent regions. The overall homology with both alleles of *Plasmodium falciparum* PMMSA is 33%.

Key words: Malaria; *Plasmodium chabaudi chabaudi*; Precursor to the major merozoite surface antigens; Nucleic acid sequence; Epitope mapping

Introduction

At the end of its intraerythrocytic development the *Plasmodium* parasite undergoes several rounds of nuclear division and forms a number of individual merozoites. After rupture of the infected red cell the merozoites are released into the bloodstream and rapidly reinvade new erythrocytes. Since merozoites represent the only stage in the erythrocytic cycle that is directly exposed to the immune system, they are considered to be important targets for vaccination [1].

The major surface antigens of the merozoite are derived from a high-molecular-weight precursor glycoprotein, that is often referred to as PMMSA (precursor to the major merozoite surface antigens). This antigen is synthesized late in the erythrocytic cycle and is subsequently processed into a number of smaller protein fragments that are associated with the surface of the mature merozoite [2–7]. The M_r of the precursor, as determined by SDS-polyacrylamide gel electrophoresis (SDS-PAGE), varies from 185–205 kDa in *Plasmodium falciparum* [8] to 230 kDa (*Plasmodium yoelii yoelii*) [9] or 250 kDa (*P. chabaudi chabaudi*) [10–11]. PMMSA shows considerable size and antigenic polymorphism between different isolates of *P. falciparum* [12] and *P. c. chabaudi* [13]. The PMMSA gene has been cloned and sequenced for a number of different isolates of *P. falciparum* [14–21]. Comparison of these sequences shows that the PMMSA gene consists of blocks that are highly conserved and blocks that vary significantly between isolates [16]. Within each variable block only two distinct sequences have been

Correspondence address: Willy Deleersnijder, Laboratorium Algemene Biologie, Instituut voor Moleculaire Biologie, Paardenstraat 65, B-1640 St. Genesius Rode, Belgium

Note: Nucleotide sequence data reported in this paper have been submitted to the GenBank™ data base with the accession number M34947.

Abbreviations: PMMSA, precursor to the major merozoite surface antigens; p199, PMMSA of *P. c. chabaudi* IP-PC1; p197, PMMSA of *P. y. yoelii* YM; mAb, monoclonal antibody; IIF, indirect immunofluorescence; AA, amino acid.

found with the exception of the variable block closest to the amino-terminus where three versions have been found [22]. Furthermore most sequences contain near the aminoterminal end a region of tripeptide tandem repeats that is highly polymorphic between different strains.

Partial sequence information is also available on the PMMSA gene of *Plasmodium vivax* [23] and the complete sequence of the *P. y. yoelii* YM PMMSA gene has recently been published [24].

Although several research groups have been able to induce partial or complete protection with purified PMMSA [25–28] or synthetic oligopeptides derived from it [29–31], the exact mechanism by which this protective immunity operates is still unclear and probably involves both humoral and cell-mediated immunity [32–35]. As a first step towards developing a mouse model system in which these questions might more easily be addressed we have cloned and sequenced the PMMSA gene of the rodent malaria parasite *P. c. chabaudi* IP-PC1 and also established a crude epitope map of PMMSA.

Materials and Methods

Parasites. Strain IP-PC1 of *P. c. chabaudi* [36], obtained from Dr. P. Falanga (Institut Pasteur Paris) was cloned by limiting dilution. One clone, termed IP-PC1/C was used for the sequence analysis described in this study. Strain IP-PC1 is a rat-adapted strain that was transferred to mice where it induces fairly synchronous infections. IP-PC1 schizont infected erythrocytes do not sequester. Several attempts to mosquito-transmit IP-PC1 or IP-PC1/C were unsuccessful. Parasites were grown in OF1 outbred mice (Iffa Credo), kept in an inverted nycthemeral cycle for diurnal schizogony.

Monoclonal antibodies. Hybridomas secreting PMMSA specific monoclonal antibodies (mAbs) 1–7, 9–10, 50 and 52 were generated in this laboratory. Spleen cells from hyperimmune BALB/c mice were fused with myeloma cell-line NSO/U [37]. Hybridoma cultures producing antiplasmodial antibodies were identified by indirect immunofluorescence (IIF) and cloned by limiting dilution. Ascites fluid from pristane (2,6,10,14 tetramethylpentadecane; Aldrich)-primed mice was used as the source of mAbs. PMMSA-specific mAbs were identified on the basis of their surface reactivity with purified merozoites in suspension (IIF) and immunoprecipitation of an approximately 250-kDa antigen. Mice were made hyperimmune by nivaquine treatment at a parasitemia of 5–25% followed by two more parasite challenges (10^7 infected red cells per mouse) at three-week intervals.

PMMSA-specific mAbs 12.3, 12.11, 12.12, 12.15, 12.17, raised against a cloned *P. c. chabaudi* isolate (isolate CB) [38] were kindly provided by D. Walliker (University of Edinburgh, U.K.). PMMSA-specific mAbs H98 and H100 were a generous gift from M. Hommel (University of Liverpool, U.K.) and had been raised against clone PC-7 of the *P. c. chabaudi* isolate IP-PC1 [39].

Preparation of cDNA and genomic libraries. Parasitized blood was collected when infection reached 30–50% and parasites were predominantly at schizont stage. Leukocytes were removed from infected blood as described elsewhere [40]. RNA was extracted by homogenization of saponin-liberated schizonts in 6 M guanidinium-HCl/0.1 M Na-acetate, pH 5.2 and centrifugation through a 4.8 M CsCl/10 mM EDTA (pH 8.0) cushion at 35 000 rev./min in a Beckman SW 41 rotor for 16 h. Poly $(A)^+$ RNA was selected by oligo(dT) cellulose chromatography and cDNA prepared according to the Amersham cDNA synthesis kit protocol. The cDNA was subsequently methylated with *Eco*RI methylase and ligated to phosphorylated *Eco*RI linkers with T4 DNA ligase. This mixture was then cleaved with *Eco*RI and fractionated on a Bio-Gel A-50m column (Biorad). Fractions containing cDNA molecules > 500 bp were ligated to dephosphorylated λgt11 *Eco*RI arms and packaged in vitro (Packagene, Promega).

For the construction of the genomic library, *Eco*RI/*Xba*I cleaved genomic DNA was ligated into dephosphorylated lambda GEM-2 *Eco*RI/*Xba*I arms (Promega) and packaged in vitro.

Screening of the libraries. Screening of recombi-

med mice was
PMMSA-specific
sis of their sur-
zoites in suspen-
on of an approx-
ere made hyper-
a parasitemia of
asite challenges
:) at three-week

12.11, 12.12.
cloned *P. c.*
8] were kindly
y of Edinburgh.
198 and H100
mel (University
raised against
isolate IP-PC1

mic libraries.
when infection
were predomi-
ytes were re-
ibed elsewhere
ogenization of
guanidinium-
centrifugation
DTA (pH 8.0)
ckman SW 41
as selected by
y and cDNA
m cDNA syn-
subsequently
and ligated to
T4 DNA lig-
d with *Eco*RI
-50m column
lA molecules
rylated λgt11
(Packagene.

genomic li-
c DNA was
bda GEM-2
packaged in

g of recombi-

nant phages with monoclonal sera was done according to Huynh et al. [41]. For hybridization screenings DNA probes were radiolabeled using the 'Multiprime DNA Labelling System' (Amersham) and hybridized to plaque blots on Hybond-N (Amersham) according to standard protocols [42]. Final washes occurred under stringent conditions (10 min at 65°C in 15 mM NaCl/1.5 mM Na-citrate).

Sequencing. Inserts of selected clones were subcloned into M13mp18 or pUC 18 sequencing vectors. Plasmid or M13 subclones were then generated that contained progressive unidirectional deletions of each insert by controlled exonuclease III digestion (Erase-a-Base, Promega). Sequencing was done on either single-stranded (M13) or double-stranded (pUC18) templates by the dideoxy chain-termination method of Sanger et al. [43] using modified T7 DNA polymerase (Sequenase, USB). Both strands were entirely sequenced in the coding and 3′ untranslated areas and partially in the 5′ untranslated region. Computer-assisted storage and analysis of sequence data was facilitated using the PC/GENE software package (Intelligenetics).

Epitope mapping. Plaque-purified recombinant cDNA λgt11 phages 72, 100, 46 and 452, expressing parts of *P. c. chabaudi* PMMSA were toothpicked from suspensions containing approximately 10^7 plaque forming units ml^{-1} onto the top agarose layer (containing approximately 10^9 *Escherichia coli* Y1090 [41] bacteria) of 90 mm diameter LB agar plates (100 μg ampicillin ml^{-1}). Plates were then incubated at 42°C for 3.5 h, subsequently overlaid with a dry nitrocellulose filter disk, saturated previously with 10 mM isopropyl β-D-thiogalactopyranoside (IPTG) in water, and incubated for an additional 3.5 h at 37°C. Nitrocellulose membranes were then saturated overnight with TBST buffer (10 mM Tris-HCl pH 8.0/150 mM NaCl, 0.05% Tween-20) containing 1% BSA and then incubated for 1 h with TBST containing mouse serum antibody or ascites fluid at a 1/200 dilution. Filters are washed several times with TBST and then incubated for 1 h with TBST containing affinity-purified alkaline phosphatase conjugated goat anti mouse IgG antibodies (dilution 1/1000). After a final round of washes substrate (0.05 mg ml^{-1} 5-bromo-6-chloro-3-indolyl acetate, 0.1 mg ml^{-1} nitro blue tetrazolium in 10 mM $MgCl_2$/100 mM NaCl/100 mM Tris-HCl, pH 9.5) is added to the filters and the enzymatic reaction is stopped after 5–15 min by rinsing the membranes with water. The relative position of cDNA expression clones 72, 100 and 46 in the open reading frame was determined by sequence analysis of the ends of their inserts. Clone 452 was positioned by restriction mapping.

Results

Cloning strategy. cDNA expression libraries (λgt11) were screened with a mixture of PMMSA-specific mAbs. Several positive clones were detected among which clones 100 and 46 were selected for sequence analysis. Clone 46 was later shown to contain a cloning artefact. This involved the fortuitous ligation of a PMMSA-specific cDNA molecule to an unrelated cDNA molecule (dotted line in Fig. 1). This conclusion was based on Southern blot analysis and comparison with the *P. y. yoelii* YM PMMSA sequence and was confirmed by sequencing genomic clone RX4. Sequence data from the nonspecific part of clone 46 were discarded for this study. Clone R16 was obtained by rescreening cDNA libraries with a radiolabeled DNA fragment that originated from the 3′ end of the clone 100 insert.

Clone 100 insert hybridized on Southern blots of *Eco*RI cleaved genomic DNA to a single band of approximately 12 kb, whereas a single 5.4-kb band was detected on Southern blots

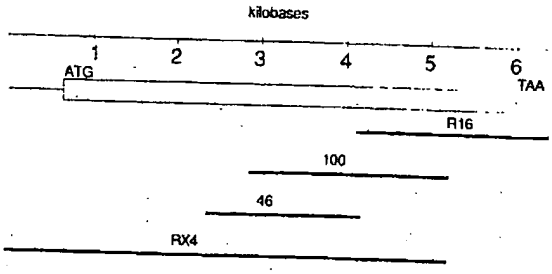


Fig. 1. Cloning strategy for the p199 gene. Clones 46, 100 and R16 are λgt11 cDNA-expression clones. Clone RX4 is a genomic λ clone. The section of clone 46 that originated from a fortuitous ligation event is represented by a dotted line.

of *Eco*RI/*Xba*I cleaved genomic DNA (data not shown). This 5.4-kb band was cloned by screening a genomic library of *Eco*RI/*Xba*I cut genomic DNA ligated into λ GEM2 vector with radiolabelled clone 100 insert. In this way genomic clone RX4 (insert size = 5.4 kb) was obtained.

Sequence analysis. The inserts of overlapping clones RX4, 46, 100 and R16 were subcloned into M13mp18 or pUC18 vectors and partially or totally sequenced. Together these 4 inserts spanned a region of 6409 bp. This sequence is presented in Fig. 2. One major open reading frame can be found, starting at nucleotide 667 and terminating with a stop codon at position 6022. This codes for a protein of 1785 amino acids (AA) with a calculated M_r of 198 886. The A+T content of the sequence is high with an average of 67% in the coding region and 83% in the 5′ and 3′ untranslated region. The A/T ratio of the coding strand is 1.74. This biased A/T ratio of mRNA sense strands appears to be a general phenomenon for malarial genes [44].

The encoded protein has many of the basic features of other PMMSA. At the aminoterminal end a putative signal peptide (residues 1–19) is present while a stretch of hydrophobic amino acids, probably functioning as a membrane anchor sequence, is found at the C-terminus (residues 1765–1785). Ten out of a total of 20 Cys residues are located in the last 110 AA of the protein. Eight potential *N*-glycosylation sites are present. Several tandem repeat oligopeptides, mostly incompletely conserved, are scattered throughout the protein. Analysis at the nucleotide level shows that the individual repeat units are clearly related. Most conspicuous is a stretch of incompletely conserved 7*6 AA starting at residue 324. Although the other tandem repeat structures are not so extensive, they do occur in many different types. In addition to the above mentioned hexapeptides, tri-, tetra-, penta- and heptapeptide tandem repeats can be observed. Also, a stretch of 7 consecutive alanine residues and a string of 7 consecutive aspartic acid residues are present in the sequence. These can be considered as monocodon repeats.

Comparison to PMMSA of other species. The *P. c. chabaudi* PMMSA sequence (p199) was aligned to known PMMSA sequences of other species using the PALIGN program (PC/GENE). With *P. y. yoelii* YM PMMSA (p197) an overall homology of 69% was detected at the AA level. This homology is not equally distributed along the protein sequence but is clustered in large zones of high homology interspersed with 4 areas of very poor homology (Fig. 3). Large insertions and/or deletions have occurred in these areas. Interestingly, all repetitious sequences are found in these regions. The 20 Cys residues present in both proteins are completely conserved. Alignment with either *P. falciparum* PMMSA allelic sequence (isolates K1 and MAD20) displayed 33% overall homology. This homology is almost exclusively confined to the conserved and semiconserved blocks, as defined by Tanabe [16] (data not shown). The few patches of high homology that occur in the variable blocks are very often also conserved between the 2 *P. falciparum* alleles. The degree of homology appears to be as high with the semi-conserved as with the conserved blocks. From this alignment it is also clear that the 4 divergent areas in the p199/P197 comparison 'coincide' with the *P. falciparum* variable blocks 4, 8, 10 and 14 (as defined by Tanabe).

A crude epitope map of p199 was established by screening the reactivity of a battery of 18 *P. c. chabaudi* PMMSA specific monoclonal antibodies with a set of overlapping cDNA expression clones in λgt11. Fig. 4 summarizes the results. The major conclusion from this analysis is that all monoclonals seem to map to the central third part of the molecule and that none is binding to the hexapeptide tandem repeats. This area does not seem to be very immunogenic. These data also indicate that carbohydrate moieties do not play a major role in the immunogenicity of the molecule.

Discussion

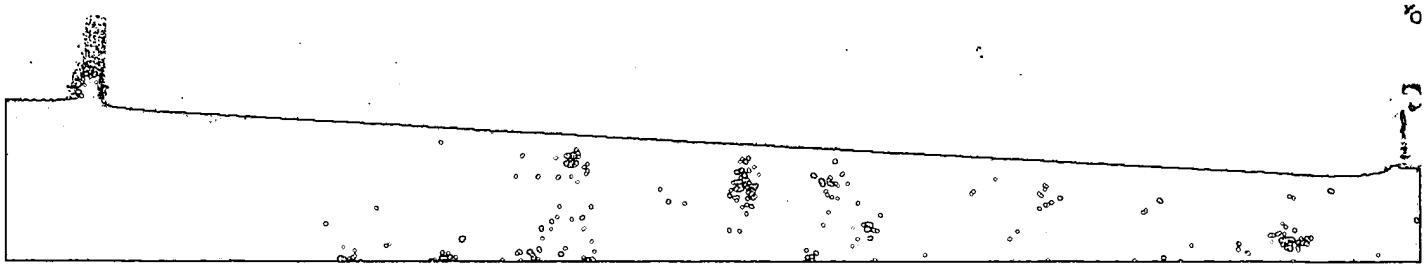
In this study we present the complete primary structure of *P. c. chabaudi* PMMSA (p199) based on the DNA sequence of the corresponding gene. P199 exhibits similar characteristics to other PMMSA. It has a calculated M_r of 199 000. shows putative signal and membrane anchor sequences and a clustering of Cys residues in the last 120 AA.

Although the predicted molecular weights for

r species us-
). With *P. y.*
ll homology
This homol-
e protein se-
of high ho-
ery poor ho-
or deletions
gly, all rep-
egions. The
ns are com-
ier *P. falci-*
ites K1 and
ology. This
ned to the
as defined
ew patches
ible blocks
te 2 *P. fal-*
y appears
s with the
it is also
199/P197
rium vari-
Tanabe).
tablished
f 18 *P.*
ntibodies
on clones
The ma-
ll mono-
urt of the
iexapep-
em to be
cate that
r role in

te pri-
(p199)
spond-
stics to
99 000.
ior se-
in the

its for



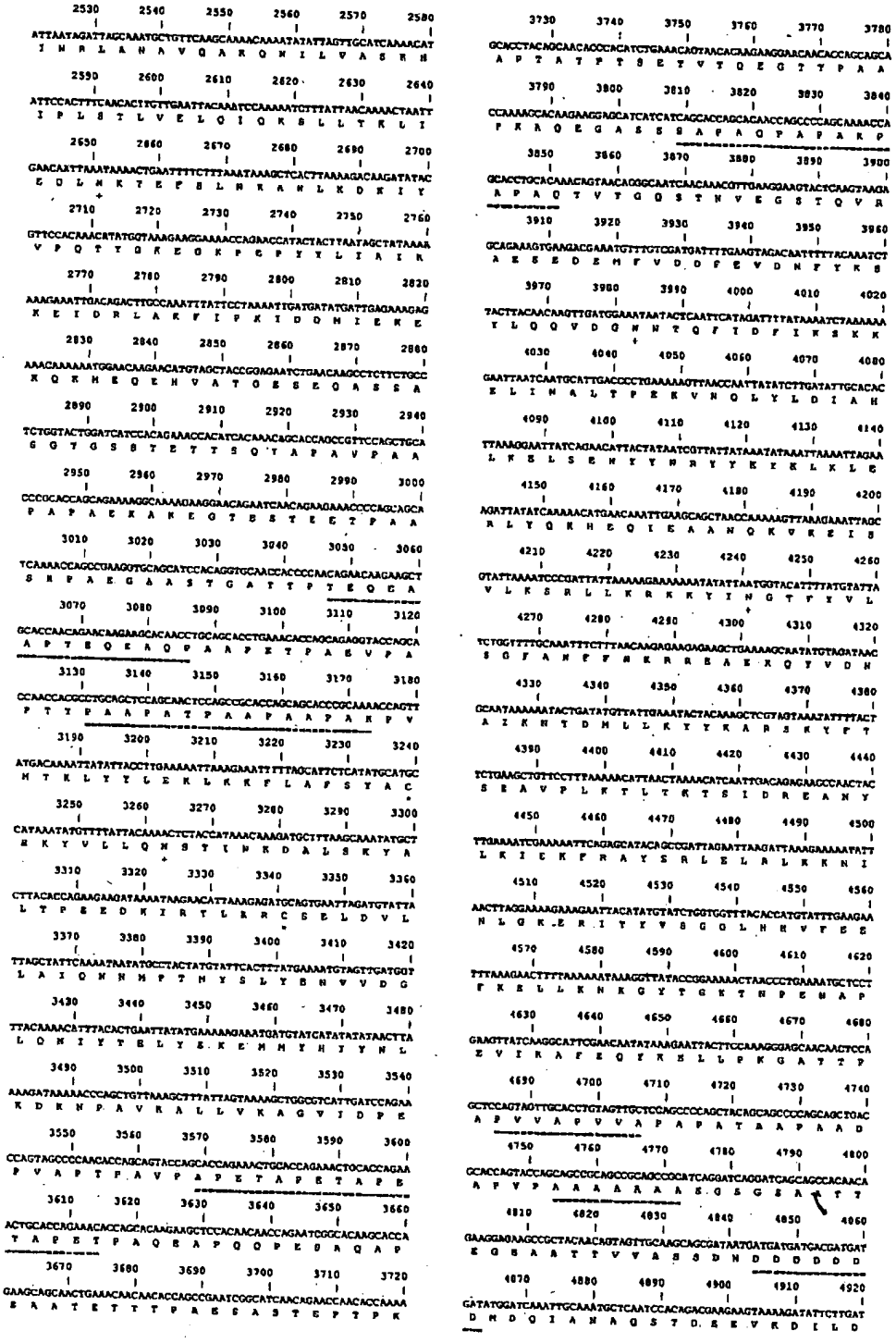


Fig. 2. Nucl
reading frame
are doubly u

PMMSA f
voelii (197
kDa) [14–
molecules
grate more
falciparum
tween appa
M_r has beer
itive antige
nounced hy
and the pres
the proper
gens are re
p199 and p
P. falciparu

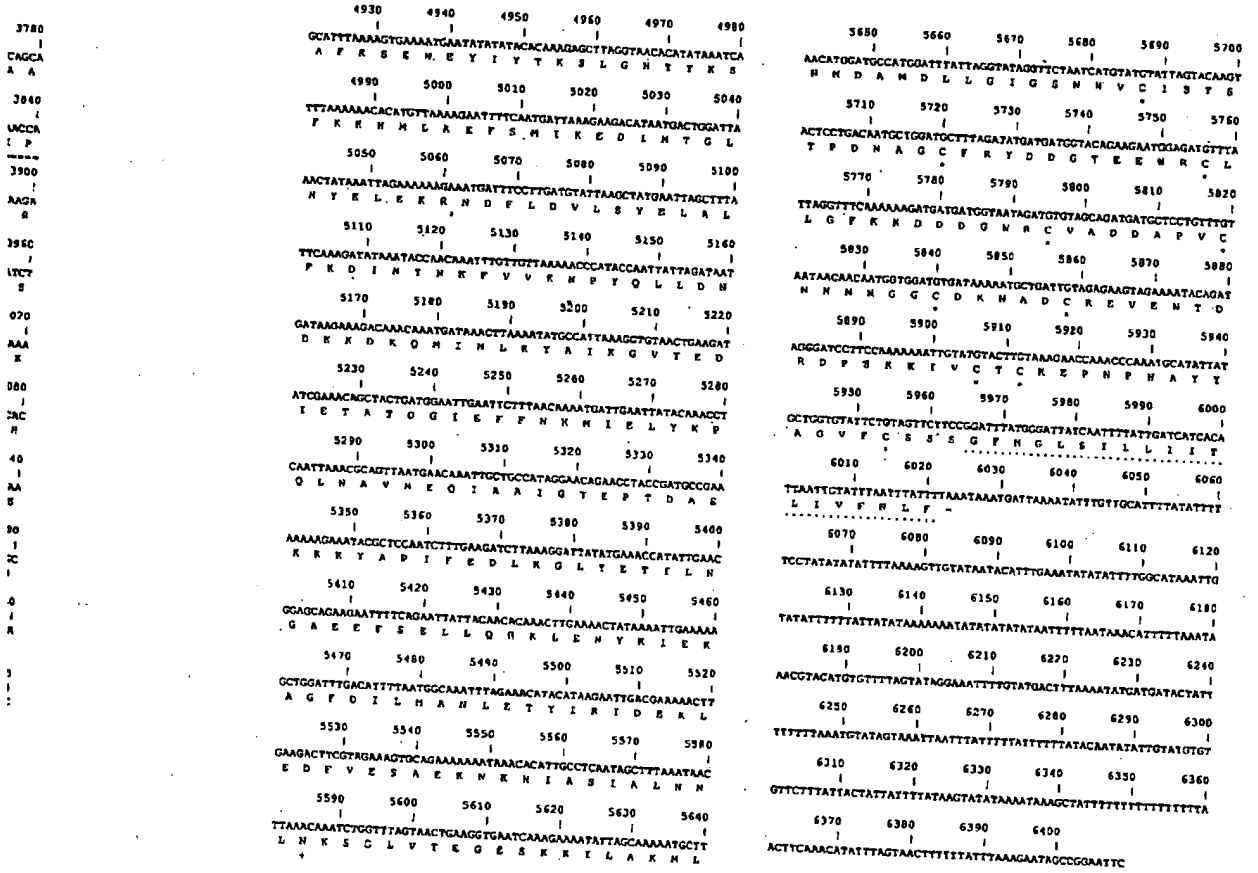


Fig. 2. Nucleotide sequence of the p199 gene of *P. c. chabaudi* and the deduced amino acid sequence of the major open reading frame. Putative signal and membrane anchor sequences are underlined by a dotted line, whilst tandem repeat structures are doubly underlined. Cys residues and potential *N*-glycosylation sites are denoted by asterisks and '+' signs respectively. Nucleotide 186 is as yet undetermined.

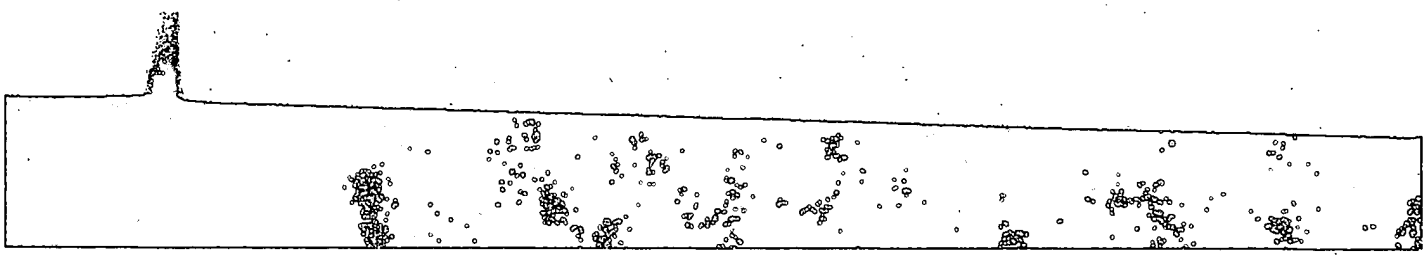
PMMSA from *P. c. chabaudi* (199 kDa), *P. y. yoelii* (197 kDa) [24] and *P. falciparum* (188–196 kDa) [14–16,18–19] are rather similar, PMMSA molecules from both rodent malaria parasites migrate more slowly on SDS-PAGE than their *P. falciparum* counterparts [45]. A discrepancy between apparent M_r on SDS-PAGE and predicted M_r has been generally observed for malarial repetitive antigens [46]. It is thought that the pronounced hydrophilic character of these antigens and the presence of repeat structures interfere with the proper binding of SDS so that these antigens are retarded in SDS-PAGE. However both p199 and p197 are neither more hydrophilic than *P. falciparum* PMMSA nor do they contain dramatically more repeats (p197 actually has less). The most plausible explanation would therefore be that PMMSA of rodent malaria parasites are more extensively glycosylated than *P. falciparum* PMMSA. This glycosylation could occur at one or more of the 4 potential *N*-glycosylation sites that are conserved between *P. c. chabaudi* and *P. y. yoelii* (residues 662, 866, 1106 and 1640 in p199). On the other hand Holder and Freeman were unable to find experimental evidence for glycosylation in *P. y. yoelii* PMMSA [9]. The physicochemical basis of this retarded migration in SDS-PAGE remains therefore unclear.

Southern blots probed with clone 100 insert only showed 1 band. Also not a single nucleotide

238

```
P199      - MKAIGLLFSFVFFAIYCKSETIGVYNDLVHKLEKLEELSVEGLELFQKSQVIVNA -55
            :: :::::.:::::: *::::: ::::: :::::: ::: :::::::::::: ::
P197      - MKVIGLLFSFVFFAIKCKSETIEVYNDLIQKLEKLESLSVDGLELFQKSQVIINA -55

P199      - QSPETPVDPFTNPEFAQKLQPFILKFEELGFTEQTELVNLIKTLGPNKYGLKYLI -110
              :    ::::: ::: : :  ::: :::::::::::::: : :: :: ::::
P197      - TQPTETIDPFTNHNFAQQVQDFVTKFEGLGFTEQTELVNLIKALTPNRYGVKYLI -110

P199      - ESKEEFNELMHAINFYYDVLRDKLNDMCANNYCEIPEHLKINVEEIEMLKKVVLG -165
            ::::::: :::::::::::::::::::*::::*::::::::  :: :::::: ::
P197      - ESKEEFNGLMHAINFYYDVLRDKLNDMCANNYCEIPEHLKISEEETEMLKKVILG -165

P199      - YRKPIENIQDDLVKLEEYIARNKATAETLNTLITEETKKITPEEETDCNDTNCDN -220
            ::::::::::::  ::: :: ::: :   :: :: :::::: ::   :*::  *:
P197      - YRKPIENIQDDIEKLEIYIERNKETVAALNALIAEETKKIQPEGNEDCNDASCDS -220

P199      - TKYGKKKAIYQAMYNVIFYKKQLAEIKKVIEVLEKRVATLKKNEAIKPLLQQIEA -275
            :: ::: :::::::::::::::::: :: :::::::: ::::: ::::: ::::
P197      - DKYNKKKPIYQAMYNVIFYKKQLAEIQKVVEVLEKRVSTLKKNDAIKPLWQQIEV -275

P199      - IRGPPAVTEG-QIATEGSSEETKQNSTESSNTKTTTTDKAVTTQTATKATGTETN -329
              : ::   :: : :      : :::                          :
P197      - LNAAPVVTAETQIVTGG------QSSTEP----------------------GSGGS -303

P199      - TGTETNTGTETNTATGTTTATGTTTATGTPTV---------TEPVQVPAVQVLTE -375
              :         ::   :      : :: :              ::  :  ::
P197      - SASGTSSSGQASAGTGVEQANTVASVTVTPSVGQNGEASTNPQTAQVQPVPTLTL -358

P199      - EEKAKKIAELYAQIKEIAKTIKFNLDGIFVDPVELEYYKKEKKNESCH-STSSCH -429
            ::: :::: :::::::::::::::: :::::: :::: ::::: ::*  ::::*
P197      - EEKQKKIAGLYAQIKEIAKTIKFNLEGIFVDPIELEYFKKEKKKESCNLSTSSCK -413

P199      - KNKTPETVIPLNVRYPNGISYPLTEEVVYSKIAHNAAETTYGDLTNVDNTAITED -484
            :::  :: :::  :::::::::: :  :: ::: :::::::::::  :::  : :
P197      - KNKASETIIPLTIRYPNGISYPLPENDVYNKIANNAAETTYGDLTHPDNTPLTGD -468

P199      - LTTNEQARKNLIKAIKKKIEAEEQKLVELKDDYDTKLAAFNGQKTPFKEAAKKFY -539
            : ::::::: ::::::::: ::: ::  ::  :: ::  :: ::::::::::: ::
P197      - LATNEQARKDLIKAIKKKIKAEEKKLETLKTNYDNKLTEFNQQKTPFKEAAKEFY -523

P199      - ESKFRNKLTTDIFDDFKTKRTEYMNKKAALVGCEYGNTQQLINKLNKQLNYLQDY -594
            :::::::::  ::  ::::: ::: ::  :  *:::::  ::::::::::::::
P197      - ESKFRNKLTSEIFEKFKTKRDEYMTKKTELNTCEYGNTKELINKLNKQLNYLQDY -578

P199      - GLRKEIVNTEIEYFSNKKSELQYNINRLANAVQAKQNILVASKHIPLSTLVELQI -649
             ::: :   ::::::::: :::::::::: :::::::: :::::  ::::::::::
P197      - SLRKDIISNEIEYFSNKKKELQYNINRLAEAVQAKQNVLVASKDVPLSTLVELQI -633

P199      - QKSLLTKLIEQLNKTEFSLNKAHLKDKIYVPQTYGKEGKPEPYYLIAIKKEIDRL -704
            ::::::: ::::+::: ::::: :::: ::: ::: :::::::::::: ::: :::
P197      - QKSLLTKQIEQLNKTEVSLNKAQLKDKLYVPKTYGNEGKPEPYYLIAVKKEVDRL -688

P199      - AKFIPKIDDMIEKEKQKMEQEHVATGESEQASSASGTGSSTETTSQTA-------- -752
            : :::::  :: :::  :::    :::::  :     :::    :
P197      - AQFIPKIESMIAKEKERMEQGPAITGESEEVPSGPSAESSTDRSTQSSTSSSSSS -743
```

```
P199   - ----------PAVPAAPAPAEKAKEGTESTEETPAASKPAEGAASTGATTPTEQ -796
                  : ::::::: :   ::  :::
P197   - SSTPAAAESSSATLPEAPAPAEAASPSTEASEETT---------------------- -778

P199   - EAAPTEQEAQPAAPETPAEVPAPTTPAAPATPAAPAAPAKPVMTKLYYLEKLKKF -851
                          : :   :   :  ::::::::::::::::: ::
P197   - ------------------IPPTTQETQPSQAASSTTPAKPVMTKLYYLEKLQKF -814

P199   - LAFSYACHKYVLLQNSTINKDALSKYALTPEEDKIRTLKRCSELDVLLAIQNNMP -906
         : ::: *::::::::+:::::::::::::: :::::::::::*::::::::::::::
P197   - LVFSYSCHKYVLLQNSTINKDALSKYALTSEEDKIRTLKRCSELDVLLAIQNNMP -869

P199   - TMYSLYENVVDGLQNIYTELYEKEMMYHIYNLKDKNPAVKALLVKAGVIDPEPVA -961
         ::::::: :::::::::::::::::::::::: ::: ::  : :::::::: :::::
P197   - TMYSLYESIVDGLQNIYTELYEKEMMYHIYKLKDENPSIKSLLVKAGVIEPEPVA -924

P199   - PTPAVPAPETAPE---TAPETAPETPAQEAPQQPESAQAPEAATETTTPAESAST -1013
           :   :     :     ::          : ::     :::    ::
P197   - APTPVTPAATEQQQQQATPDVQSDAPAP-----SDVSQQPETPVTSTTPEVTTST -974

P199   - EPTPKAP----------------------TATPTSETVTQEGTTPAAPKAQEGASS -1047
         :   ::                        :::          :
P197   - EASSSAPGEGTPSGEAGASGTEGATASNAATPAGTSASGSAASNASTTSDVTPPA -1029

P199   - SAPAQPAPAKPAPAQ--TVTGQSTNVEGSTQVRAESEDEMFVDDFEVDNFYKSYL -1100
           : : :    :::::      :: :       :::::  :  :::: :: :::::
P197   - AAAAVPSTSTPAPAQPPAANSQSGNPDSGIRSRAESEEDMPADDFELDNLYKSYL -1084

P199   - QQVDGNNTQFIDFIKSKKELINALTPEKVNQLYLDIAHLKELSEHYYNRYYKYKL -1155
         :: ::+:: :: ::::::::: :::::::::::: :::::::::::: ::  :::
P197   - QQIDGNNTEFINFIKSKKELIKALTPEKVNQLYLEIAHLKELSEHYYDRYSTYKL -1139

P199   - KLERLYQKHEQIEAANQKVKEISVLKSRLLKRKKYINGTFYVLSGFANFFNKRRE -1210
         :::::: :::::   :      : :: ::::::   :: :: : :  ::::::::
P197   - KLERLYNKHEQIQLTNRQIRDLSILKARLLKRKQTLNGVFYILNGYVNFFNKRRE -1194

P199   - AEKQYVDNAIKNTDMLLKYYKARSKYFTSEAVPLKTLTKTSIDREANYLKIEKFR -1265
         ::::::::: :::::::::::: :::::::::::: : : ::: :::::::::
P197   - AEKQYVDNALKNTDMLLKYYKARTKYFTSEAVPLKTLSKASLDRESNYLKIEKFR -1249

P199   - AYSRLELRLKKNINLGKERITYVSGGLHHVFEEFKELLKNKGYTGKTNPENAPEV -1320
         :::::::::::::::::::: ::::::::::::::::: : : :::: :: :::::
P197   - AYSRLELRLKKNINLGKERISYVSGGLHHVFEEFKELIKDKDYTGKKNPDNAPEV -1304

P199   - IKAFEQYKELLPKGATTPAPVVA---PVVAPAPATAAPAADAPVPAAAAAAASGS -1372
         :::::::::::: :   : ::      : ::::   :    :: :   : :
P197   - TNAFEQYKELLPKGVTVSTPAVAVTTTLAADAPATPEGAVPGAVPGAVPGAVPGA -1359

P199   - GSAATTEGEAATTVVASSDNDDDDDDDMDQIANAQSTDEEVKDILDAFKSENEYI -1427
           :       : :  ::  ::  :::  :::  :: :   :::: :  :  :
P197   - VPGAVPGSGTDTRVAGSSV-DDNEDDDIYQIASGQSEDAPEKDILSEFTNESLYV -1413

P199   - YTKSLGNTYKSFKKHMLKEFSMIKEDIMTGLNYKLEKRNDFLDVLSYELALFKDI -1482
         ::: :: :::: ::::: ::: ::::   ::: :  :::::: ::: :: ::::
P197   - YTKRLGSTYKSLKKHMLREFSTIKEDMTNGLNNKSQKRNDFLEVLSHELDLFKDL -1468
```



```
P199      - NTNKFVVKNPYQLLDNDKKDKQMINLKYAIKGVTEDIETATDGIEFFNKMIELYK -1537
P197      - STNKYVIRNPYQLLDNDKKDKQIVNLKYATKGINEDIETTTDGIKFFNKMVELYN -1523

P199      - PQLNAVNEQIAAIGTEPTDA---EKKKYAPIFEDLKGLYETILNGAEEFSELLQH -1589
P197      - TQLAAVKEQIATIEAETNDTNKEEKKKYIPILEDLKGLYETVIGQAEEYSEELQN -1578

P199      - KLENYKIEKAGFDILMANLETYIRIDEKLEDFVESAEKNKHIASIALNNLNKSGL -1644
P197      - RLDNYKNEKAEFEILTKNLEKYIQIDEKLDEFVEHAENNKHIASIALNNLNKSGL -1633

P199      - VTEGESKKILAKMLNMDAMDLLGIGSNHVCISTS-TPDNAGCFRYDDGTEEWRCL -1698
P197      - VGEGESKKILAKMLNMDGMDLLGVDPKHVCVDTRDIPKNAGCFRDDNGTEEWRCL -1688

P199      - LGFKKDDDGNRCVADDAPVCNNNNGGCDKNADCREVENTDRDPSKKIVCTCKEPN -1753
P197      - LGYKKGE-GNTCVENNNPTCDINNGGCDPTASCQNAESTEN--SKKIICTCKEPT -1740

P199      - PNAYYAGVFCSSSGFMGLSILLIITLIVFNLF -1785
P197      - PNAYYEGVFCSSSSFMGLSILLIITLIVFNIF -1772
```

Fig. 3. Alignment of p199 with *P. yoelii* PMMSA (p197). Residues that are part of repetitive structures in both p199 and p197 are in bold face. Positions of Cys residues that are conserved are indicated by asterisks whilst conserved potential *N*-glycosylation sites are denoted by '+' signs. The four highly divergent areas have been boxed.

of difference was observed between genomic clone RX4 and cDNA clones 100 and 46 in the areas where both types of clone have been sequenced (totalling about 1500 nucleotides). Taken together these data provide strong evidence that the p199 gene occurs as a single copy in the genome.

The protein sequence of p199 is very homologous to p197 (69%) and 33% homologous to both *P. falciparum* PMMSA allelic sequences. The same level of homology (31%) had previously been observed between these 2 allelic sequences and p197 [24], indicating that, as might have been expected, *P. c. chabaudi* and *P. y. yoelii* are evolutionary equidistant to *P. falciparum*. The major findings of the PMMSA interspecific comparisons, as described here, are schematically represented in Fig. 5 and can be summarized as follows: (i) p199 and p197 are rather homologous but differ extensively in 4 areas that correspond to *P. falciparum* PMMSA variable blocks 4, 8, 10 and 12. On the contrary variable blocks 2, 6 and 16 are well conserved between p199 and p197. (ii) The homology that exists between p199/p197 and *P. falciparum* PMMSA is situated almost exclusively in the conserved and semi-conserved blocks. The variable blocks are nearly totally divergent. In these variable areas small patches of homology occur that frequently correspond to stretches that are also conserved between the 2 *P. falciparum* alleles.

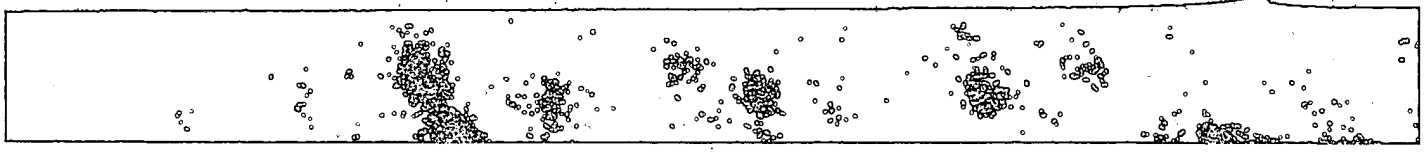
It is assumed that the *P. falciparum* alleles evolved in 2 biologically isolated populations that later on merged [16]. The rodent malaria sequence data further illustrate the evolutionary behavior of PMMSA: some areas of the protein are well conserved whereas others are very variable. It is interesting that whereas 7 hypervariable areas are apparent when comparing the 2 *P. falciparum* alleles or when comparing *P. falciparum* PMMSA with the rodent malaria PMMSA, *P. c. chabaudi* and *P. y. yoelii* only diverge profoundly in 4 of these 7 variable blocks. This might indicate that variable blocks 4, 8,10 and 14 evolve even faster than variable blocks 2, 6 and 16.

The question might be asked as to why some regions in the molecule did evolve very rapidly while other parts changed at a slower pace. Clearly there must be structural and/or functional constraints on the more conserved parts of the protein. These data might however also indicate that the

Fig. 4. Crude
reading frame
mAbs that rea
is indicated by

divergent zc
which incre
types of pos
are the need
and immun
considered t
invasion of

Fig. 5. Schem
with the PMM
into conserved
where p199 ar
and MAD20/K
homology to
been kept unc
occurred in so



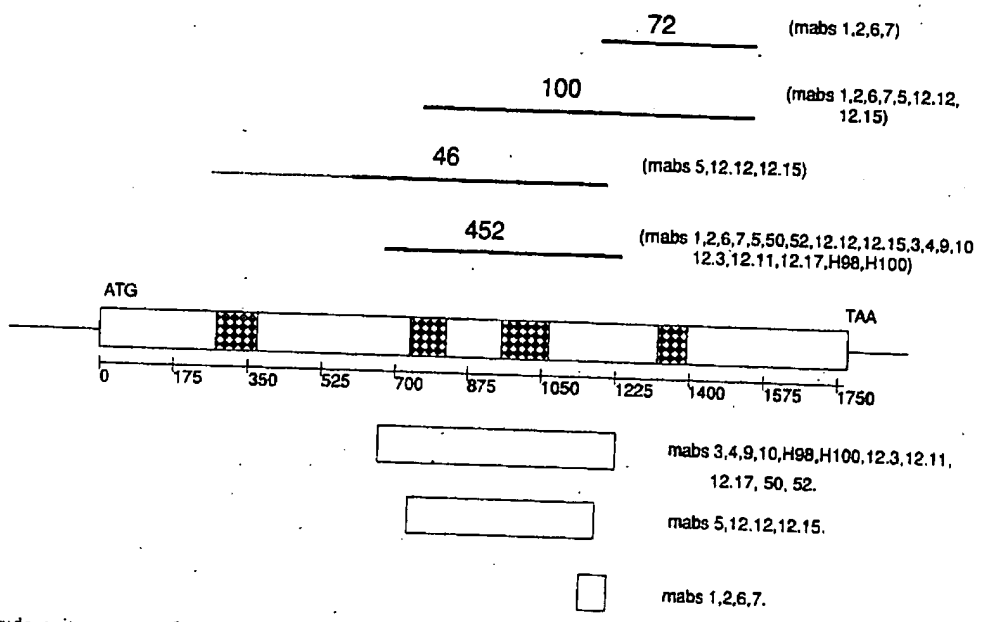


Fig. 4. Crude epitope map of p199. The relative positions of λgt11 expression clones 72, 100, 46 and 452 to the total open reading frame are indicated. Amino acid numbering is indicated below the bar representing the open reading frame. Names of mAbs that react with each of these expression clones are bracketed. The section of clone 46 that represents the cloning artefact is indicated by a dotted line. The 4 divergent regions in p199 are denoted by chequered blocks. The areas to which various epitopes can be mapped are indicated by open boxes.

divergent zones were subject to positive selection which increased the rate of genetic change. Two types of positive selection that might be envisaged are the need to adapt to an evolving vertebrate host and immunological pressure. Since PMMSA is considered to participate in the recognition and/or invasion of red cells [47], the variable blocks, which differ considerably between different *Plasmodium* species and between the 2 *P. falciparum* alleles, might constitute the domains that mediate this interaction. It must be assumed then that the 2 *P. falciparum* PMMSA alleles evolved to interact with different structures on the human red cell membrane (as proposed by Tanabe et al. [16]) and

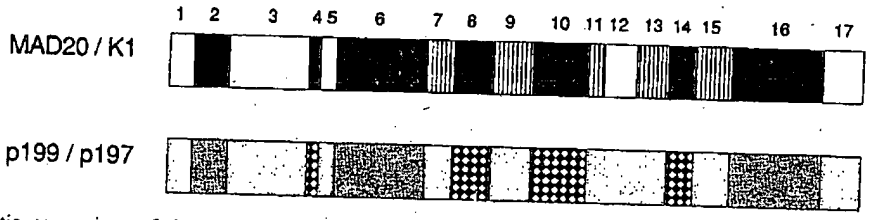


Fig. 5. Schematic comparison of the amino acid sequences of the 2 *P. falciparum* PMMSA alleles (strains MAD20 and K1) with the PMMSA amino acid sequences from *P. c. chabaudi* (p199) and *P. y. yoelii* (p197). *P. falciparum* sequences are divided into conserved blocks (open boxes), semi-conserved blocks (boxes with vertical lines) and variable blocks (black boxes). Areas where p199 and p197 are highly divergent are denoted by chequered boxes. Regions that show homology between p199/p197 and MAD20/K1 are indicated by stippled boxes whilst blocks that are highly conserved between p199 and p197 but show no homology to MAD20/K1 are shaded. For clarity the relative sizes of the different blocks in the *P. falciparum* alleles have been kept unchanged in the p199/p197 block diagram. It should be noted however that due to insertions/deletions that have occurred in some variable blocks of p199/p197 the relative size of these blocks is not accurately represented. The numbering of the different blocks is according to Tanabe et al. [16].

also that *P. c. chabaudi* and *P. y. yoelii* PMMSA recognize different groups on the mouse red cell. It is however possible that the need to interact with different red cell structures in different hosts can be accommodated by minor AA changes in the more conserved areas.

Alternatively the rapid genetic change in the divergent areas might have been generated by immunological pressure. The repetitive motifs which are found in these areas might be the result of special genetic mechanisms that warrant rapid diversification for the evasion of immune responses. The strain-specific protection shown by PMMSA of *P. c. chabaudi* AS and CB [48] and the isolation of anti-PMMSA mAb resistant lines from cloned *P. c. chabaudi* AS [49] is in keeping with this hypothesis. Interesting also in this respect is the observation that *P. c. chabaudi* PMMSA appears to be even more polymorphic than *P. falciparum* PMMSA since every field isolate out of 15 tested belonged to a different PMMSA serotype (McLean, A.P., Ph.D.-Thesis, University of Edinburgh, 1986).

At the same time these variable regions are probably not very immunodominant. This is indicated by our epitope mapping studies. The first variable area (block 4) is not recognized by any of the PMMSA specific monoclonals while the epitopes of 4 mAbs were shown to map outside of the 4 divergent areas. On the basis of its similar structure (alternating tripeptide repeats) the major repeat area in p199 (residues 324–365) appears to be the homologue to the tripeptide repeats seen in PMMSA of most *P. falciparum* isolates. However the *P. falciparum* tripeptide tandem repeats occur in variable block 2 whereas the p199 repeats are found in variable block 4. The tripeptide tandem repeat area in *P. falciparum* PMMSA is widely polymorphic among different isolates. Work is in progress to determine whether a similar polymorphism prevails in different *P. c. chabaudi* strains.

Acknowledgements

This work was supported by a Smith Kline RIT/IWONL contract and by EEC contracts TSD-M-152 and TS-2-0148. We wish to thank Drs. M. Hommel and D. Walliker for providing mAbs and Drs. Jef Seurinck and Gaston Mathijssen (Plant Genetic Systems) for help with sequencing work.

References

1 Newbold, C.I. (1985) Parasite antigens in protection diagnosis and escape: *Plasmodium*. In: Current Topics in Microbiology and Immunology, pp. 69–104. Springer-Verlag, Berlin, Heidelberg.

2 Hall, R., Osland, A., Hyde, J.E., Simmons, D.L., Hope, I.A. and Scaife, J. (1984) Processing, polymorphism, and biological significance of p190, a major surface antigen of the erythrocytic forms of *Plasmodium falciparum*. Mol Biochem. Parasitol. 11, 61–80.

3 Lyon, J.A., Geller, R.H., Haynes, J.D. and Weber, J.L. (1986) Epitope map and processing scheme for the 195 000-dalton surface glycoprotein of *Plasmodium falciparum* merozoites deduced from cloned overlapping segments of the gene. Proc. Natl. Acad. Sci. USA. 83, 2989–2993.

4 Freeman, R.R. and Holder, A.A. (1983) Surface antigens of malaria merozoites. A high molecular weight precursor is processed to an 83,000-mol.wt. form expressed on the surface of *Plasmodium falciparum* merozoites. J. Exp. Med. 158, 1647–1653.

5 David, P.H., Hadley, T.J., Aikawa, M. and Miller, L.H. (1984) Processing of a major parasite surface glycoprotein during the ultimate stages of differentiation in *Plasmodium knowlesi*. Mol. Biochem. Parasitol. 11, 267–282.

6 Heidrich, H-G., Matzner, M., Miettinen-Baumann, A. and Strych, W. (1986) Immunoelectron microscopy shows that the 80 000-dalton antigen of *Plasmodium falciparum* merozoites is localised in the surface coat. Z. Parasitenkd. 72, 681–683.

7 Atkinson, C.T., Aikawa, M., Fujino, T., Tam, L.Q., Hui, G.S.N. and Siddiqui, W.A. (1987) Ultrastructural localization of protective and non-protective *Plasmodium falciparum* proteins using serum samples from vaccinated *Aotus* monkeys. J. Parasitol. 73, 1235–1240.

8 Holder, A.A. (1988) The precursor to major merozoite surface antigens: structure and role in immunity. Prog. Allergy. 41, 72–97.

9 Holder, A.A. and Freeman, R.R. (1984) Characterisation of a high molecular weight protective antigen of *Plasmodium yoelii*. Parasitology. 88, 211–219.

10 Boyle, D.B., Newbold, C.I., Smith, C.C. and Brown, K.N. (1982) Monoclonal antibodies that protect in vivo against *Plasmodium chabaudi* recognize a 250,000-dalton parasitic polypeptide. Infect. Immun. 38, 94–102.

11 Newbold, C.I., Boyle, D.B., Smith, C.C. and Brown, K.N. (1982) Identification of a schizont and species specific surface glycoprotein on erythrocytes infected with rodent malarias. Mol. Biochem. Parasitol. 5, 45–54.

12 McBride, J.S., Newbold, C.I. and Anand R. (1985) Polymorphism of a high molecular weight schizont antigen of the human malaria parasite *Plasmodium falciparum*. J. Exp. Med. 161, 160–179.

13 Newbold, C.I., Schryer, M., Boyle, D.B., McBride, J.S., McLean, A., Wilson, R.J.M. and Brown, K.N. (1984) A possible molecular basis for strain specific immunity to malaria. Mol. Biochem. Parasitol. 11, 337–347.

14 Holder, A.A., Lockyer, M.J., Odink, K.G., Sandhu, J.S., Riveros-Moreno, V., Nicholls, S.C., Hillman, Y., Davey, L.S., Tizard, M.L.V., Schwarz, R.T. and Freeman, R.R. (1985) Primary structure of the precursor to the three major

surface antigens of *Plasmodium falciparum* merozoites. Nature. 317, 270–273.
15 Mackay, M., Goman, M., Bone, N., Hyde, J.E., Scaife, J., Certa, U., Stunnenberg, H. and Bujard, H. (1985) Polymorphism of the precursor for the major surface antigens of *Plasmodium falciparum* merozoites: studies at the genetic level. EMBO J. 4, 13B, 3823–3829.
16 Tanabe, K., Mackay, M., Goman, M. and Scaife, J.G. (1987) Allelic dimorphism in a surface antigen gene of the malaria parasite *Plasmodium falciparum* J. Mol. Biol. 195, 273–287.
17 Weber, J.L., Leininger, W.M. and Lyon, J.A. (1986) Variation in the gene encoding a major merozoite surface antigen of the human malaria parasite *Plasmodium falciparum*. Nucleic Acids Res. 14, 3311–3323.
18 Peterson, M.G., Coppel, R.L., McIntyre, P., Langford, C.J., Woodrow, G., Brown, G.V., Anders, R.F. and Kemp, D. (1988) Variation in the precursor to the major merozoite surface antigens of *Plasmodium falciparum*. Mol. Biochem. Parasitol. 27, 291–302.
19 Chang, S.P., Kramer, K.J., Yamaga, K.M., Kato, A., Case, S.E. and Siddiqui, W.A. (1988) *Plasmodium falciparum*: gene structure and hydropathy profile of the major merozoite surface antigen (gp 195) of the Uganda-Palo Alto isolate. Exp. Parasitol. 67, 1–11.
20 Myler, P.J. (1989) Nucleotide and deduced amino acid sequence of the gp195 (MSA-1) gene from *Plasmodium falciparum* Palo Alto PLF-3/B11. Nucleic Acids Res. 17, 5401.
21 Certa, U., Rotmann, D., Matile, H. and Reber-Liske, R. (1987) A naturally occurring gene encoding the major surface antigen precursor of p190 of *Plasmodium falciparum* lacks tripeptide repeats. EMBO J. 6, 4137–4142.
22 Peterson, M.G., Coppel, R.L., Moloney, M.B. and Kemp, D.J. (1988) Third form of the precursor to the major merozoite surface antigens of *Plasmodium falciparum*. Mol. Cell. Biol. 8, 2664–2667.
23 Del Portillo, H.A., Gysin, J., Mattei, D.M., Khouri, E., Udagama, P.V., Mendis, K.N. and David, P.H. (1988) *Plasmodium vivax*: cloning and expression of a major blood-stage surface antigen. Exp. Parasitol. 67, 346–353.
24 Lewis, A.P. (1989) Cloning and analysis of the gene encoding the 230-kilodalton merozoite surface antigen of *Plasmodium yoelii*. Mol. Biochem. Parasitol. 36, 271–282.
25 Holder, A.A. and Freeman, R.R. (1981) Immunization against blood-stage rodent malaria using purified parasite antigens. Nature 294, 361–364.
26 Siddiqui, W.A., Tam, L.Q., Kramer, K.J., Hui, G.S.N., Case, S.E., Yamaga, K.M., Chang, S.P., Chan, E.B.T. and Kan, S-C. (1987) Merozoite surface coat precursor protein completely protects Aotus monkeys against *Plasmodium falciparum* malaria. Proc. Natl. Acad. Sci. USA 84, 3014–3018.
27 Hall, R., Hyde, J.E., Goman, M., Simmons, D.L., Hope, I.A., Mackay, M. and Scaife, J. (1984) Major surface antigen gene of a human malaria parasite cloned and expressed in bacteria. Nature. 311, 379–382.
28 Perrin, L.H., Merkli, B., Loche, M., Chizzolini, C., Smart, J. and Richle, R. (1984) Antimalarial immunity in Saimiri monkeys. Immunization with surface components of asexual bloodstages. J. Exp. Med. 160, 441–451.
29 Cheung, A., Leban, J., Shaw, R.A., Merkli, B., Stocker, J., Chizzolini, C., Sander, C. and Perrin, L.H. (1986) Immunization with synthetic peptides of a *Plasmodium falciparum* surface antigen induces antimerozoite antibodies. Proc. Natl. Acad. Sci. USA. 83, 8328–8332.
30 Patarroyo, M.E., Romero, P., Torres, M.L., Clavijo, P., Moreno, A., Martinez, A., Rodriguez, R., Guzman, F. and Cabezas, E. (1987) Induction of protective immunity against experimental infection with malaria using synthetic peptides. Nature. 328, 629–632.
31 Patarroyo, M.E., Amador, R., Clavijo, P., Moreno, A. Guzman, F., Romero, P., Tascon, R., Franco, A., Murillo, L.A., Ponton, G. and Trujillo, G. (1988) A synthetic vaccine protects humans against challenge with asexual blood stages of *Plasmodium falciparum* malaria. Nature. 332, 158–161.
32 Freeman, R.R. and Holder, A.A. (1983) Characteristics of the protective response of Balb/c mice immunized with a purified *Plasmodium yoelii* schizont antigen. Clin. Exp. Immunol. 54. 609–616.
33 Playfair, J.H.L., De Souza, J.B., Freeman, R.R. and Holder, A.A. (1985) Vaccination with a purified blood-stage malaria antigen in mice: correlation of protection with T cell-mediated immunity. Clin. Exp. Immunol. 62, 19–23.
34 Majarian, W.R., Daly, T.M., Weidanz, W.P. and Long, C.A. (1984) Passive immunization against murine malaria with an IgG3 monoclonal antibody. J. Immunol. 132, 3131–3137.
35 Burns, J.M., Jr., Daly, T.M., Vaidya, A.B. and Long, C.A. (1988) The 3′ portion of the gene for a *Plasmodium yoelii* merozoite surface antigen encodes the epitope recognized by a protective monoclonal antibody. Proc. Natl. Acad. Sci. USA. 85, 602–606.
36 David, P.H., Hommel, M., Benichou, J.C., Eisen, H. and Pereira da Silva, L.H. (1978) Isolation of malaria merozoites: Release of *Plasmodium chabaudi* merozoites from schizonts bound to immobilized concanavalin A. Proc. Natl. Acad. Sci. USA 75, 5081–5084.
37 Kearney, J.F., Radbruch, A., Liesegang, B. and Rajewsky, K. (1979) A new mouse myeloma cell line that has lost immunoglobulin expression but permits the construction of antibody-secreting hybrid cell lines. J. Immunol. 123, 1548–1550.
38 Killick-Kendrick, R. and Peters, W. (1978) Rodent Malaria, Academic Press, London, New York. San Francisco.
39 Hommel, M., David, P.H., Guillotte, H. and Pereira da Silva, L. (1982) Protection against *Plasmodium chabaudi* malaria, I. Vaccination of mice with merozoites and Freund's adjuvants. Ann. Immunol. (Inst. Pasteur) 133C. 57–67.
40 Howard, R.J., Smith, P.M. and Mitchell, G.F. (1978) Removal of leucocytes from red cells in *Plasmodium berghei*-infected mouse blood and purification of schizont-infected cells. Ann. Trop. Med. Parasitol. 72, 573–575.
41 Huynh, T.V., Young, R.A. and Davis, R.W. (1986) Constructing and screening cDNA libraries in lambda gt10 and lambda gt11. In: DNA Cloning Vol. I (Glover, D.M., ed.), pp. 49–78, IRL Press, Oxford/Washington DC.
42 Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning. A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
43 Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA. 74, 5463–5467.
44 Weber, J.L. (1988) Molecular biology of malaria parasites. Exp. Parasitol. 66, 143–170.
45 Holder, A.A., Freeman, R.R. and Newbold, C.I. (1983) Serological cross reaction between high molecular weight proteins synthesised in blood schizonts of *Plasmodium yoelii*, *Plasmodium chabaudi* and *Plasmodium falciparum*.

Mol. Biochem. Parasitol. 9, 191–196.
46 Anders, R.F., Coppel, R.L., Brown, G.V. and Kemp, D.J. (1988) Antigens with repeated amino acid sequences from the asexual blood stages of *Plasmodium falciparum*. Prog. Allergy. 41, 148–172.
47 Hadley, T.J., Klotz, F.W. and Miller, L.H. (1986) Invasion of erythrocytes by malaria parasites: a cellular and molecular overview. Annu. Rev. Microbiol. 40, 451–477.
48 Bates, M.D., Newbold, C.I., Jarra, W. and Brown, K.N. (1988) Protective immunity to malaria: studies with cloned lines of *Plasmodium chabaudi chabaudi* in CBA/Ca mice. III. Protective and suppressive responses induced by immunization with purified antigens. Parasite Immunol. 10, 1–15.
49 Wood, J.C., Sales de Aguiar, J.C., Jarra, W., Ogun, S.A., Snounou, G. and Brown, K.N. (1989) In vivo selection of populations of *Plasmodium chabaudi chabaudi* AS resistant to a monoclonal antibody that reacts with the precursor to the major merozoite surface antigen. Infect. Immun. 57, 2128–2135.

Molecular and
Elsevier
MOLBIO 014

S
Schi

(

Centre d

Molecular n
characterizatio
(Sm39) and mc
invertebrate de
parasite and in
The mollusc se
species. The h
could be speci
the vesicles pr
around sporoc

Key words: *Sc*

Introductio

Adaptatio
developmen
gies to eva
molecular r
of these po
tosomes sha
invertebrate
non-specific
as from the
[3,4]. In thi
port the im
antigenic m

Correspondenc
logie et de Bi
CNRS 624. In

Note: Nucleot
been submitte
number M 36

Abbreviations:
SDS, sodium

0166-6851/90/

Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 4030–4034, May 1991
Medical Sciences

Primary structure of the merozoite surface antigen 1 of *Plasmodium vivax* reveals sequences conserved between different *Plasmodium* species

(parasitic protozoa/malaria/vaccine/gene cloning)

HERNANDO A. DEL PORTILLO*, SHIRLEY LONGACRE, ELISABETH KHOURI, AND PETER H. DAVID

Unite d'Immunoparasitologie, Institut Pasteur, 25 rue du Dr. Roux, Paris 75015, France

Communicated by William Trager, January 9, 1991 (received for review December 17, 1990)

ABSTRACT **Merozoite surface antigen 1 (MSA1) of several species of plasmodia has been shown to be a promising candidate for a vaccine directed against the asexual blood stages of malaria. We report the cloning and characterization of the MSA1 gene of the human malaria parasite *Plasmodium vivax*. This gene, which we call *Pv200*, encodes a polypeptide of 1726 amino acids and displays features described for MSA1 genes of other species, such as signal peptide and anchoring sequences, conserved cysteine residues, number of potential N-glycosylation sites, and repeats consisting here of 23 glutamine residues in a row. When the nucleotide and deduced amino acid sequences of the MSA1 of *P. vivax* are compared to those of another human malaria parasite, *Plasmodium falciparum*, and to those of the rodent parasite *Plasmodium yoelii*, 10 regions of high amino acid similarity are observed despite the very different dG+dC contents of the corresponding genes. All of the interspecies conserved regions reside within the conserved or semiconserved blocks delimited by the sequences of different alleles of the MSA1 gene of *P. falciparum*.**

The surface of the invasive merozoite of plasmodia constitutes one of the potential targets of a vaccine directed against the blood stages of malaria. Merozoite surface antigen 1 (MSA1), described by Holder and Freeman in 1982 (1), has been extensively studied in the human malarial parasite *Plasmodium falciparum* (reviewed in ref. 2). There are several allelic forms of this polymorphic high molecular weight antigen, and conserved, semiconserved, and variable regions can be found in the different alleles (3–5). The antigen is processed on the surface of the merozoite, although the exact stage at which processing occurs is subject to discussion (6). MSA1 has also been shown to bind in a specific manner to the surface of erythrocytes and could thus constitute one of the merozoite surface ligands involved in invasion of the erythrocyte (7).

A number of immunization experiments performed with parasite-derived or recombinant MSA1 or with MSA1 peptides in monkeys (reviewed in ref. 2) as well as in humans (8) point to this antigen as one of the most promising vaccine candidates against malaria asexual blood stages. *P. falciparum* is the only human malarial parasite for which the protective properties of the MSA1 have been assessed. Since protective immunity in malaria is species-specific (9), it is unlikely that a vaccine against one species will protect against others. Although *Plasmodium vivax* is the most widely distributed human malaria parasite, little is known about the properties of MSA1 in this species (10); this is partly due to the difficulty in obtaining large quantities of a parasite that cannot be maintained in continuous culture. The cloning and characterization of the gene coding for the MSA1 of *P. vivax* should allow appropriate immunization studies to be performed with recombinant proteins.

A portion of the *P. vivax* MSA1 gene (Belem strain) has been previously characterized (11), and we present here the complete primary structure of this gene,† which we call *Pv200*. The organization of *Pv200* is similar to that of the MSA1 gene of *P. falciparum*, *Pf190* (3, 12), and to that of the rodent malaria parasite *Plasmodium yoelii*, *Py230* (13). There are 10 regions of high amino acid similarity conserved among the three parasite species. Since this molecule, like many other *P. vivax* antigens, is otherwise polymorphic (14, 15), such regions of interspecies conservation could be of importance in the development of an asexual stage malaria vaccine.

MATERIALS AND METHODS

Parasites. The *P. vivax* Belem strain, adapted to *Saimiri* monkeys, was used for the production of DNA (11).

Construction and Screening of Genomic DNA Libraries. Two DNA libraries were constructed: (*i*) Library A. Genomic DNA was completely digested with *Eco*RI and 5 μg was fractionated on a 1% agarose gel. Fragments between 5 and 15 kilobases (kb) were electroeluted from a slice of the gel, extracted with phenol, and precipitated with ethanol. Pellets were washed, dried, and dissolved in double-distilled H_2O. A 1-μg aliquot was ligated into the *Eco*RI arms of the λ vector gt*WES* (GIBCO/BRL) according to the supplier's instructions. The library was obtained by transforming LE392 competent cells and it was screened with a 1.9-kb DNA insert containing a portion of the *Pv200* gene, Pv200/1.9 (see *Results*) (11).

(*ii*) Library B. A 0.5-μg sample of *Hin*dIII-digested DNA was ligated into the *Hin*dIII site of the vector pBR322 treated with calf intestinal alkaline phosphatase (Pharmacia) and the library was obtained by transformation of DH5 α competent cells. The library was screened with a 0.98-kb DNA insert corresponding to the first 0.98 kb from the 5′ end of the Pv200/1.9 clone (see *Results*).

All enzyme digestions and DNA manipulations were performed as recommended in Sambrook *et al.* (16).

DNA Sequences. Dideoxy chain termination sequences (17) were obtained by the production of exonuclease III overlapping deletion clones (18) or by the use of oligonucleotides (17-mers) synthesized on an Applied Biosystems PCR-Mate apparatus. Both DNA strands were sequenced for all the results presented here. Sequences were aligned and analyzed

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: MSA1, merozoite surface antigen 1; ICB, interspecies conserved block; CB, conserved block.

*Present address: Departamento de Parasitologia, Instituto de Ciencias Biologicas II-Universidade de São Paulo-Cidade Universitaria, Avenida Professor Lineu Prestes 1374, São Paulo, CEP 05508, Brazil.

†The sequence reported in this paper has been deposited in the GenBank data base (accession no. M60807).

by using the DNA program of Staden (19). The sequences used for homology studies were those of *P. falciparum* MAD20 (3) and of *P. yoelii* YM (13).

RESULTS

Isolation of the Genomic Clones Containing the Entire *Pv200* Gene. We have previously reported the isolation of a clone containing a 1.9-kb genomic DNA insert, clone Pv200/1.9, including a portion of the *P. vivax* Belem strain MSA1 gene (*Pv200*) (11). Using the Pv200/1.9 DNA insert, we isolated two new clones, Pv200/7.0 and Pv200/9.0, containing the 5′ and 3′ ends of the *Pv200* gene, respectively.

On Southern blots of *P. vivax* genomic DNA digested with *Eco*RI the Pv200/1.9 insert hybridized with a single 9-kb DNA fragment (not shown). Accordingly, Pv200/1.9 was used to screen 5×10^4 phage plaques from library A, and a positive clone, Pv200/9.0, was isolated. *Eco*RI digestion of DNA from this clone released a 9-kb insert, which was subcloned in the *Eco*RI site of the Bluescript vector (Stratagene). The nucleotide sequence of Pv200/9.0 showed that it contained the remaining 3.5 kb of the 3′ end of the *Pv200* gene.

A 0.98-kb fragment of Pv200/1.9 insert, obtained through digestion of Pv200/1.9 with *Hin*dIII, hybridized with a single 7-kb band on Southern blots of genomic DNA digested with *Hin*dIII (not shown). Library B was screened with the 0.98-kb fragment. A positive clone, Pv200/7.0, was isolated and shown to contain a 7-kb insert, from which the sequence of the 5′ end of the *Pv200* gene was determined.

Nucleotide Sequence of the *P. vivax* Belem Strain *Pv200* Gene. The complete nucleotide and deduced amino acid sequences of the *Pv200* gene are shown in Fig. 1. A methionine start codon at base 91 initiates a single open reading frame of 5178 bases that finishes with the first TAA stop codon at base 5259. An A+T-rich noncoding region follows after this stop codon. Three observations indicate that the methionine codon at position 91 is the initiation codon *in vivo*. (*i*) There are two stop codons immediately upstream, at positions 64 and 76. (*ii*) A poly(A) sequence precedes this ATG, possibly representing the consensus sequence for translation initiation as described for several plasmodial genes (20). (*iii*) The amino acid sequence immediately following this ATG codon has all the features of a putative signal peptide (21). The sequence presented here is based entirely on genomic DNA fragments. We believe, however, that the *Pv200* gene contains no introns, since a continuous open

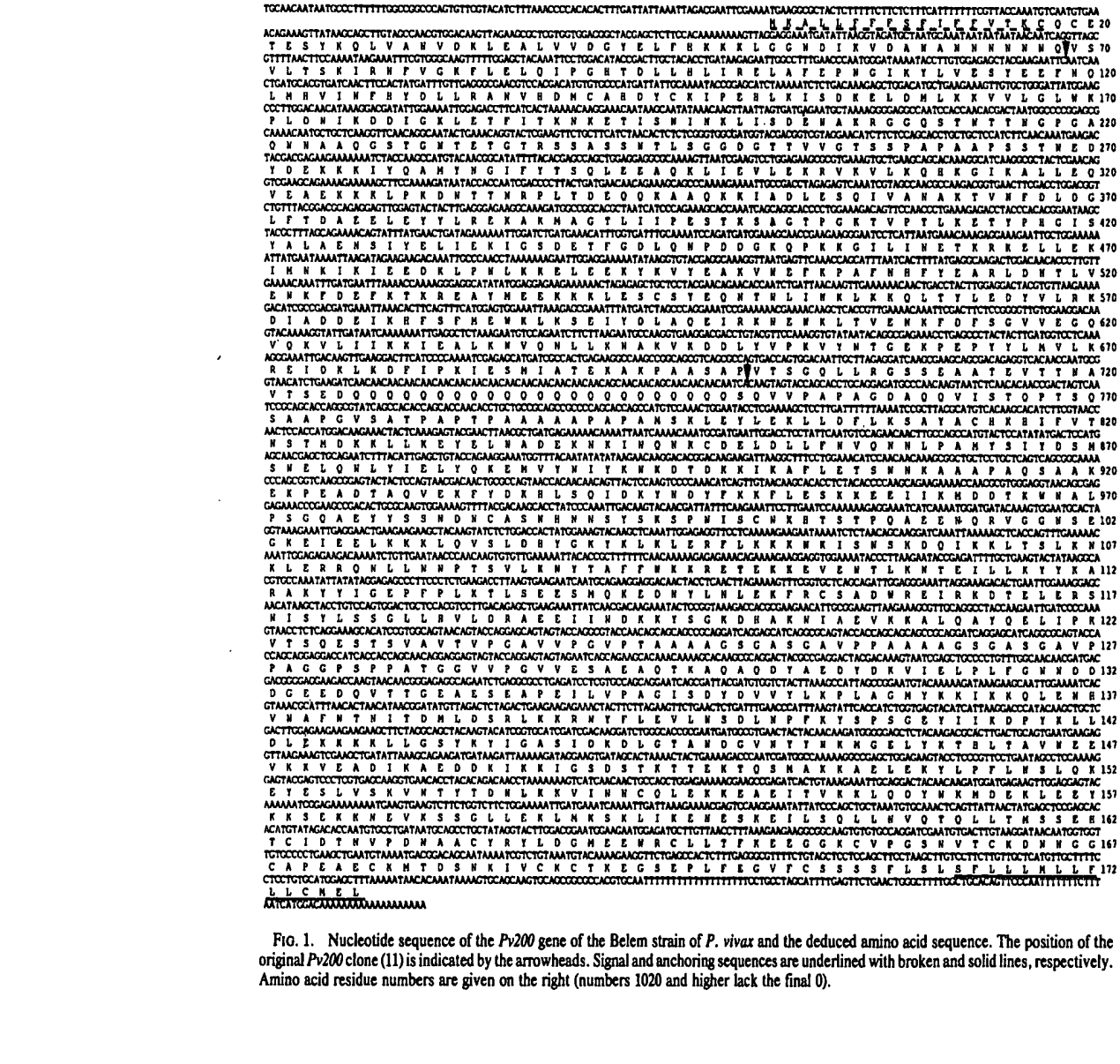


FIG. 1. Nucleotide sequence of the *Pv200* gene of the Belem strain of *P. vivax* and the deduced amino acid sequence. The position of the original *Pv200* clone (11) is indicated by the arrowheads. Signal and anchoring sequences are underlined with broken and solid lines, respectively. Amino acid residue numbers are given on the right (numbers 1020 and higher lack the final 0).

reading frame of 1726 amino acids with a calculated molecular weight of 194,267 is contained within the genomic fragments. This is in agreement with the absence of introns in the genes coding for the MSA1 of other species.

There are a potential signal peptide and a hydrophobic membrane anchor sequence at residues 1–17 and 1710–1726, respectively. Furthermore, there are 12 potential N-glycosylation sites (Asn-Xaa-Thr/Ser) and 22 cysteines, 11 of which are located within the last 110 residues of the COO terminus of the molecule. The Pv200 sequence also contains a stretch of 23 glutamines at residues 726–748.

Comparison of the codon usage in the MSA1 genes of *P. vivax*, *P. falciparum*, and *P. yoelii* revealed that codons which have G or C in the third position are more frequent in *P. vivax*. Consequently, the dG+dC content of the *Pv200* coding region is 43.4% and differs significantly from the dG+dC content of the coding regions of the MSA1 genes from *P. falciparum* (25.7%) and *P. yoelii* (31%).

Comparisons of the Pv200, Pf190, and Py230 Sequences. The deduced amino acid sequence from the *Pv200* gene was computer-aligned with the sequences of the Pf190 (allele *MAD20*) (Fig. 2) and Py230 YM (Fig. 3) polypeptides. There is an overall identity of 35.6% and 34.3% with the *P. falciparum* and *P. yoelii* sequences, respectively.

Interestingly, 17 out of the 22 cysteines of the Pv200 polypeptide were located at similar positions with respect to the Pf190 and Py230 sequences. These similarities include the 11 and 10 cysteines found at the COO terminus of Pf190 and Py230, respectively. In contrast, of 12 (Pv200), 15 (Pf190), and 11 (Py230) potential N-glycosylation sites, only 3 were conserved at the same positions between the *P. vivax* and the *P. falciparum* sequences, whereas only 1 was conserved between the *P. vivax* and *P. yoelii* sequences.

To determine the regions with an amino acid identity near 50% among the three parasite species, we combined the comparisons which had been made between Pv200–Pf190/Pv200–Py230 (this work) and Py230–Pf190 (13). Fig. 4 shows the result of such analysis. Seven ICBs were observed: ICB1, ICB2, ICB4, ICB5, ICB6, ICB8, and ICB10. Similarly, three other blocks (CB3, CB7, and CB9) were conserved between Pv200 and Pf190 but not between Pv200 and Py230 and thus could not be treated as bona fide ICBs. All these blocks reside within the conserved or semiconserved blocks of the Pf190 alleles (3).

DISCUSSION

We report the complete primary structure of the MSA1 gene of the *P. vivax* Belem strain, *Pv200*. The general structure of the gene resembles that of the MSA1 genes described for *P. falciparum* and *P. yoelii*, with a number of homologous regions and other features such as (*i*) conserved cysteine residues at the COO-terminal region, (*ii*) number of potential N-glycosylation sites, and (*iii*) the presence of 23 glutamines in a row, which could correspond in *P. vivax* to the repeated sequences described in the MSA1 genes of other species.

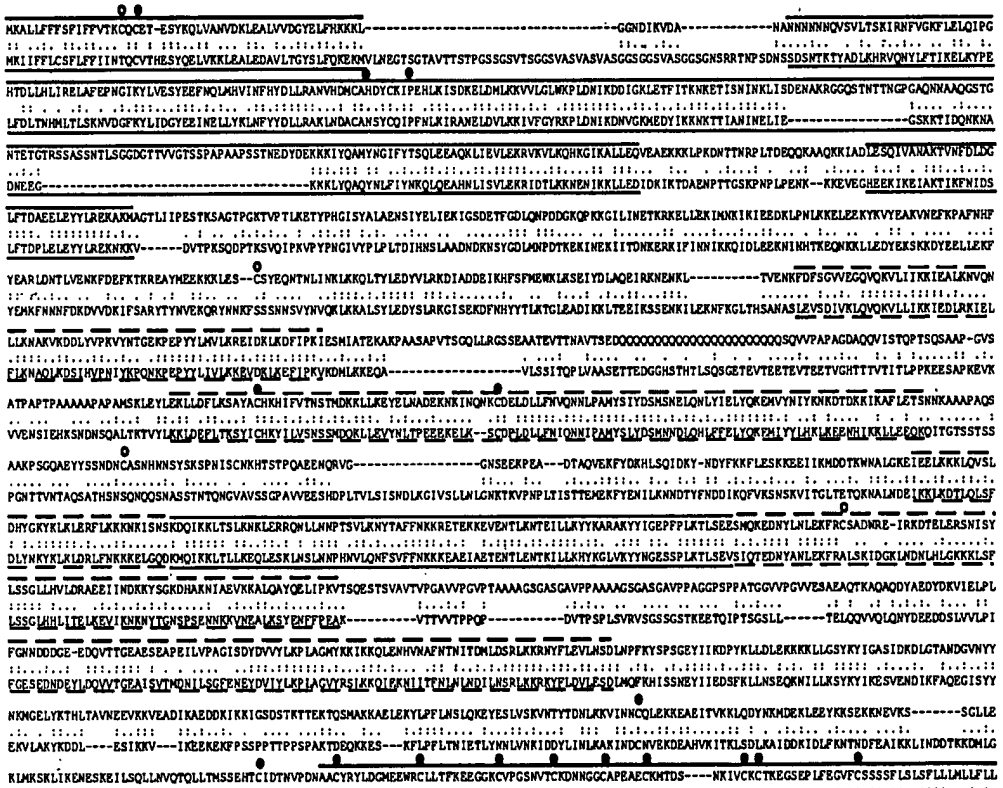


FIG. 2. Comparison of the amino acid sequences of the *P. vivax* Belem strain Pv200 (upper sequence) and of the *P. falciparum MAD20* Pf190 (lower sequence) (3). Sequences were aligned by using the program of Staden (19). Hyphens indicate gaps introduced for alignment; colons, identical residues; and periods, similar residues. Positions of the Pv200 cysteine residues conserved between these two proteins are denoted by ● and those that are not conserved, by ○. The position of the Pf190 blocks determined by the sequence of different alleles (3) is also shown; conserved blocks are indicated by unbroken underlines and overlines, semiconserved blocks are indicated by broken underlines and overlines, and variable blocks are unmarked.

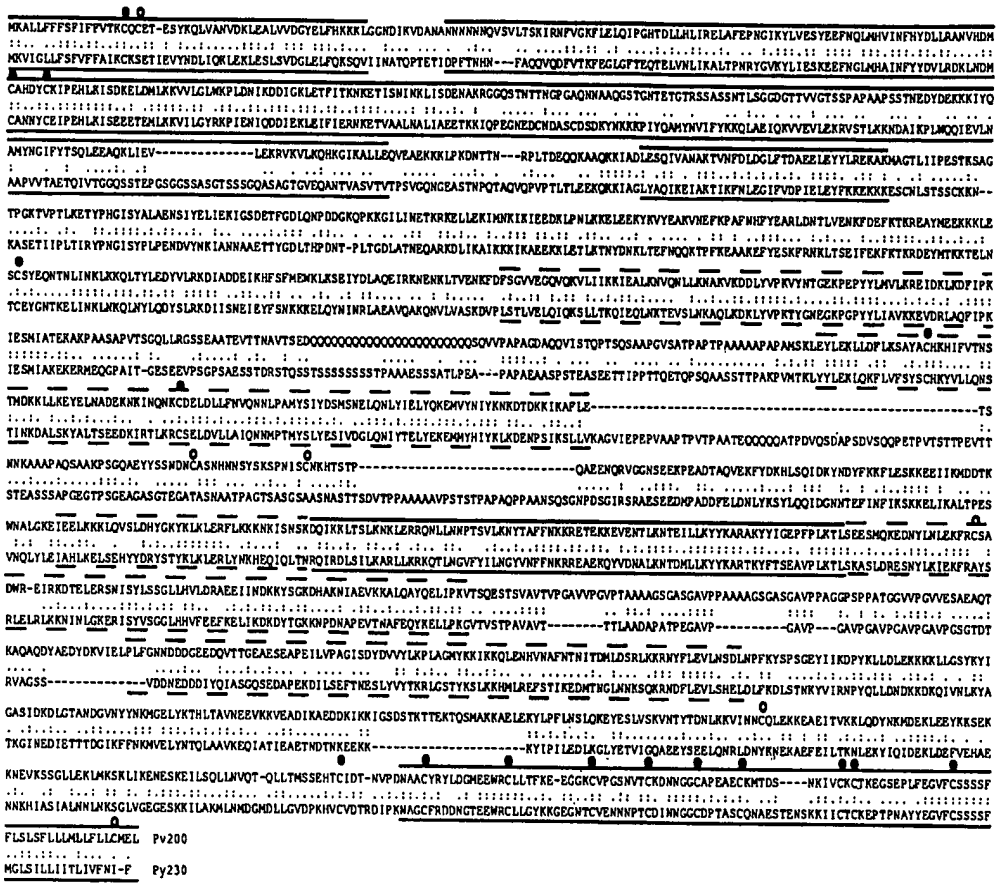


FIG. 3. Comparison of the amino acid sequences of the *P. vivax* Belem strain Pv200 (upper sequence) and of the *P. yoelii* YM Py230 (lower sequence) (13). Sequences were aligned by using the program of Staden (19). Conventions are as in Fig. 2.

Malaria parasites have been divided evolutionarily into three groups according to the base composition of their DNA (22). One group, comprising avian, rodent, and falciparum malarias, presents a genome with a low dG+dC content (18%). Another, comprising the two monkey malarias *Plasmodium knowlesi* and *Plasmodium fragile*, presents a genome with a higher dG+dC content (30%). Finally, the group of *P. vivax* and *Plasmodium cynomolgi*, human and monkey malarias which cause relapses, has a genome presenting both low and high dG+dC components. This division implies that homologous genes and their proteins should be more similar within a group than between groups (22). Our observations show that in the case of the MSA1 genes and their proteins this prediction is supported only at the nucleotide level. Indeed, the low dG+dC content of the *Pf190* and *Py230* genes leads to a higher similarity, at the nucleotide level, between them than with *Pv200*. However, when the amino acid composition is considered, Pv200 and Pf190 antigens show higher similarity and the overall distribution of their shared amino acids is more highly conserved than when Pf190 and Py230 are compared. That a higher amino acid similarity and closer overall distribution are observed in the Pv200 and Pf190 antigens despite their very different total dG+dC content most likely reflects the effects of positive selection within the human host. Accordingly, three regions of homology between the Pv200 and Pf190 antigens not conserved between the Pv200 and the Py230 antigens can be found (Fig. 4).

The analysis of the primary structure from different alleles of the MSA1 gene of *P. falciparum* allowed the definition of conserved, semiconserved, and variable regions within the molecule (3). One of the regions of amino acid identity higher than 45% conserved between the Pf190 and Py230 antigens resides within a variable block of one of the Pf190 alleles and,

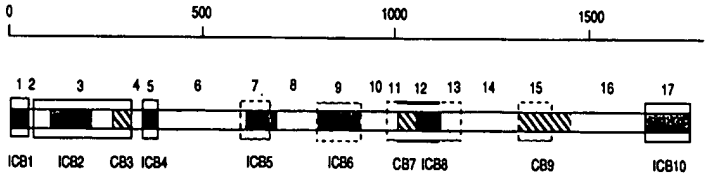


FIG. 4. Representation of the MSA1 antigen based upon amino acid conservation among the Pv200, Pf190, and Py230 proteins (inner blocks) and upon Pf190 alleles (outer blocks; solid-outline blocks, conserved areas; broken-outline blocks, semiconserved areas) (7). Shaded boxes represent interspecies conserved blocks (ICBs) with greater than 48% identity among the three parasite species. Hatched boxes represent conserved blocks (CBs) with greater than 50% identity between Pv200 and Pf190 but not between Pv200 and Py230. Open boxes represent areas of less than 45% identity. Positions of ICBs and CBs (amino acid residues of the Pv200 sequence): ICB1, 1–50; ICB2, 107–200; CB3, 274–319; ICB4, 348–387; ICB5, 620–691; ICB6, 796–895; CB7, 1040–1088; ICB8, 1092–1153; CB9, 1347–1464; and ICB10, 1622–1727.

consequently, Lewis (13) proposed the delimitation of new conserved blocks within the MSA1 antigen based on interspecies conservation. We decided to conduct a similar analysis; regions of 50 or more contiguous amino acids presenting 50% or higher identity among the three species [Pv200 vs. Pf190 and Pv200 vs. Py230 (this work) and Py230 vs. Pf190 (13)] are referred to as ICBs. Subsequently, the position of such ICBs with respect to the blocks delimited by sequences from different Pf190 alleles (3) was also examined.

All of the ICBs of MSA1 described here reside within the conserved or semiconserved blocks delimited by different alleles of the *P. falciparum* gene (Fig. 4). That such well-defined regions of MSA1 have been conserved among these three different malaria species could be explained because they are functionally or structurally important for the molecule, or because they are not immunogenic, or, finally, because immune responses against them do not block parasite growth (23). On the basis of these results, we predict that as sequences from other alleles of the *Pv200* gene are described, the general structure of the *Pv200* gene will comprise blocks that will be organized in a fashion similar to that of the blocks delimited by different *Pf190* alleles.

As for the protective properties of MSA1, most immunization trials with *P. falciparum* have used either the whole molecule or fragments from the NH_2-terminal part (reviewed in ref. 2). In particular, the two peptides used in human vaccine trials belong to the regions we have defined as ICB1 and CB3 (8). This does not exclude other portions of MSA1; in particular, ICB10 corresponds to the most COO-terminal part of the molecule. The most remarkable aspect of this part of MSA1 is that it contains more than half of all the cysteine residues that are conserved in position among the three parasite species. Significantly, a protective monoclonal antibody against a discontinuous epitope of the *P. yoelii* MSA1 has been mapped to this region (24). Immunization trials with the MSA1 antigen of *P. vivax* have yet to be reported, and the potential protective properties of Pv200 can only be extrapolated from experiments performed in other malarial species. The availability of the complete primary structure from the MSA1 gene of *P. vivax* should now allow the assessment of Pv200 as a vaccine candidate.

We are grateful to Prof. Ruth Nussenzweig, Dr. Valerie Snewin, and Dr. Alexandra Levitt for critically reviewing the manuscript; to Ms. Christiane Duponchel for providing some of the oligonucleotides used for sequencing; to Dr. Michel Rabinovitch and Dr. Jurg Gysin for support; to Dr. Denise Mattei and Dr. Artur Scherf for helpful scientific discussions; and to Dr. Mario L. A. Bressan, Dr. Charlie Roth, and Dr. Michel Keller for computer assistance. This investigation received support from the United Nations Development Programme/World Bank/World Health Organization Special Programme for Research and Training in Tropical Diseases (TDR) and from the Rockefeller Foundation.

1. Holder, A. A. & Freeman, R. R. (1982) *J. Exp. Med.* **156**, 1528–1538.
2. Holder, A. A. (1988) *Prog. Allergy* **41**, 72–97.
3. Tanabe, K., Mackay, M., Goman, M. & Scaife, J. G. (1987) *J. Mol. Biol.* **196**, 273–287.
4. Certa, U., Rotmann, D., Matile, H. & Reberliske, R. (1987) *EMBO J.* **6**, 4137–4142.
5. Kimura, E., Mattei, D. M., di Santi, S. M. & Scherf, A. (1990) *Gene* **91**, 57–62.
6. Pirson, P. J. & Perkins, M. E. (1985) *J. Immunol.* **134**, 1946–1951.
7. Perkins, M. E. & Rocco, L. J. (1988) *J. Immunol.* **141**, 3190–3196.
8. Patarroyo, M. E., Amador, R., Clavijo, P., Moreno, A., Guzman, F., Romero, P., Tascon, R., Franco, A., Murillo, L. A., Ponton, G. & Trujillo, G. (1988) *Nature (London)* **332**, 158–161.
9. Becker, F. T. (1949) in *Malariology*, ed. Boyd, M. F. (Saunders, Philadelphia), Vol. 2, pp. 1145–1157.
10. David, P. H., Barnwell, J. W. & Mendis, K. N. (1990) in *New Generation Vaccines*, eds. Woodrow, G. C. & Levine, M. M. (Dekker, New York), pp. 531–544.
11. del Portillo, H. A., Gysin, J., Mattei, D. M., Khouri, E., Udagama, P. V., Mendis, K. N. & David, P. H. (1988) *Exp. Parasitol.* **67**, 346–353.
12. Holder, A. A., Lockyer, J., Odink, K. G., Sadhu, J. S., Riveros-Moreno, V., Nicholls, S. C., Hillman, Y., Davey, L. S., Tizard, M. L.-V., Schwart, R. T. & Freeman, R. R. (1985) *Nature (London)* **317**, 270–273.
13. Lewis, A. P. (1989) *Mol. Biochem. Parasitol.* **36**, 271–282.
14. Udagama, P. V., David, P. H., Peiris, J. S. M., Ariyaratne, Y. G., Perera, K. L. R. L. & Mendis, K. N. (1987) *Infect. Immun.* **55**, 2604–2611.
15. Udagama, P. V., Gamage-Mendis, A. C., David, P. H., Peiris, J. S. M., Perera, K. L. R. L., Mendis, K. N. & Carter, R. (1989) *Am. J. Trop. Med. Hyg.* **42**, 104–110.
16. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, New York), 2nd Ed.
17. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
18. Henikoff, S. (1984) *Gene* **28**, 351–359.
19. Staden, R. (1982) *Nucleic Acids Res.* **10**, 2951–2961.
20. Saul, A. & Battistutta, D. (1990) *Mol. Biochem. Parasitol.* **42**, 55–62.
21. Mackay, M., Goman, M., Bone, N., Hyde, J. E., Scaife, J., Certa, U., Stunnenberg, H. & Bujard, H. (1985) *EMBO J.* **4**, 3823–3829.
22. McCutchan, T. F., Dame, J. B., Miller, L. H. & Barnwell, J. (1984) *Science* **225**, 808–811.
23. Weber, J. L., Leininger, W. M. & Lyon, J. A. (1986) *Nucleic Acids Res.* **14**, 3311–3323.
24. Burns, J. M., Jr., Majarian, W. R., Young, J. F., Daly, T. M. & Long, C. A. (1989) *J. Immunol.* **143**, 2670–2676.

Molecular and Biochemical Parasitology, 50 (1992) 325–334

MOLBIO 01671

Structure and expression of the gene for Pv200, a major blood-stage surface antigen of *Plasmodium vivax*

Helen L. Gibson[1], Jeffrey E. Tucker[1], David C. Kaslow[2], Antoniana U. Krettli[2], William E. Collins[3], Michael C. Kiefer[1], Ian C. Bathurst[1], and Philip J. Barr[1]

[1]*Chiron Corporation, Emeryville, CA, U.S.A.;* [2]*The Laboratory of Parasitic Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, U.S.A.; and* [3]*Division of Parasitic Diseases, Center for Infectious Diseases, Centers for Disease Control, Public Health Service, U.S. Department of Health and Human Services, Atlanta, GA, U.S.A.*

(Received 9 August 1991; accepted 21 September 1991)

Molecular cloning and structure analysis of the gene encoding the Pv200 protein of the Sal-1 strain of *Plasmodium vivax* revealed an overall identity of 34–37% when the deduced amino acid sequence was compared with the sequences of various major merozoite surface antigens of *Plasmodium falciparum*, *Plasmodium yoelii* and *Plasmodium chabaudi*. When the Sal-1 Pv200 sequence was compared with the corresponding sequence from the Belém strain of *P. vivax*, it was found that the two merozoite surface antigens were relatively well conserved with an overall amino acid sequence identity of 81%. A region of 23 repeated glutamine residues, found in the sequence of the Belém isolate was not found, however, in the Sal-1 sequence. Amino- and carboxy-terminal domains of the Pv200 protein were expressed in the yeast *Saccharomyces cerevisiae*. Each recombinant protein was shown to react with antibodies in sera from splenectomized Bolivian *Saimiri* monkeys that had been infected previously with *P. vivax*, and in human sera from individuals with a history of exposure to vivax malaria. The availability of recombinant DNA-derived Pv200 proteins will now allow a full assessment of their utility in the diagnosis and immunoprophylaxis of the benign tertian malaria associated with *P. vivax* infection.

Key words: *Plasmodium vivax*; Blood-stage antigen: Pv200; Malaria vaccine

Introduction

Large scale in vitro parasite culture cannot provide sufficient quantities of either organisms or antigens to produce a malaria vaccine. Accordingly, peptide synthesis and recombinant DNA methodologies are being evaluated extensively for the production of potential protective immunogens against the protozoan parasites that cause malaria. Because of the severity of the disease caused by *Plasmodium falciparum*, this species has represented the major target for such studies. Of the numerous potential subunit vaccine candidates from this organism, an antigen of special interest has been the major merozoite surface antigen Pf195. This antigen has been shown to reside on the surface of the schizont and, in a processed form, on the surface of the merozoite [1]. Vaccination studies with Pf195, isolated from cultured parasites, have led to high levels of protection against *P. falciparum* challenge in monkey model systems [2]. More recently, recombinant DNA-derived Pf195 antigens have been produced and studied as candidate vaccines against falciparum malaria [3–5].

Correspondence address: Philip J. Barr, Molecular Biology, Chiron Corporation, 4560 Horton Street, Emeryville, CA 94608, USA.

Note: Nucleotide sequence data reported in this paper have been submitted to the GenBank™ data base with the accession number M75674.

Abbreviations: ADH2, alcohol dehydrogenase-2; CS, circumsporozoite; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; PAGE, polyacrylamide gel electrophoresis; PCR, polymerase chain reaction; TMD, transmembrane domain.

Although less virulent than *P. falciparum*, *Plasmodium vivax* is the causative agent of benign tertian fever, a form of malaria characterized by frequent and protracted relapses. Vaccination studies against this species have been less extensive than with *P. falciparum* and have been focused almost exclusively on the sporozoite lifecycle stage of the organism [6–9]. Recently, however, a polymorphic 200-kDa component of the *P. vivax* schizont surface was defined by monoclonal antibodies, and a partial genomic clone that encoded a portion of the antigen was isolated, structurally defined and expressed in bacteria [10]. Homology of this DNA fragment, and the more recently isolated full length gene sequence [11], with the Pf195 gene, together with immunolocalization studies using antisera to the expressed protein have suggested strongly that the Pv200 protein is functionally analogous to Pf195 [10]. Here we report the molecular cloning and structure analysis of the gene for Pv200 from the Sal-1 strain [8] of *P. vivax*. We also demonstrate that amino- and carboxy-terminal domains of the protein produced in yeast can be used to detect antibodies both in monkeys and in humans previously infected with *P. vivax*.

Materials and Methods

Construction of P. vivax genomic DNA libraries. Genomic libraries were prepared as follows. 500 ng of *P. vivax* genomic DNA, from the Sal-1 strain, was digested with *Eco*RI and ligated into *Eco*RI digested λZAPII (Stratagene), packaged and introduced into *Escherichia coli* strain PLK-17. A similar *Hin*dIII digest, partially filled with dCTP and dTTP, was ligated into *Xba*I digested λZAPII that had been similarly filled with dATP and dGTP. The ligated DNA was packaged and transformed as above. Libraries of 5×10^7 and 4.1×10^7 independent clones were obtained, respectively.

Screening of P. vivax DNA. Two overlapping oligomers (45-mers), based on the sequence of del Portillo et al. [10], were labeled by the oligomer primed extension method [12], hybridized in 40% formamide-containing buffer [13] at 37°C to a Southern blot of *Eco*RI-digested *P. vivax* DNA, and washed at 65°C in 2 × SSC/0.1% SDS. A 9.5-kb *Eco*RI fragment hybridized to this probe. Two overlapping oligomers (42- and 43-mers) based on the 5′-end of the *Eco*RI clone were used to probe a Southern blot of *Hin*dIII-digested *P. vivax* DNA in a similar fashion and hybridized to a 7.0-kb fragment. Similar hybridizations were carried out on library filters and plaque purified tertiary positives were excised using *E. coli* strain XL1-Blue (Stratagene) and plasmid DNA was retransformed into *E. coli* strain D1210 for further plasmid manipulations.

Subcloning and DNA sequencing. Plasmid DNA was isolated by the alkaline lysis method [13]. Overlapping restriction fragments were subcloned into M13 vectors and both strands were sequenced by the chain-termination method [14] using M13 primers as well as specific internal primers. DNA manipulations were essentially as described [13].

Expression of amino- and carboxy-terminal domains of Pv200 in Saccharomyces cerevisiae. The polymerase chain reaction (PCR) [15] was used to amplify DNA fragments from cloned Pv200 gene sequences. Appropriate restriction sites, and in-frame initiation and termination codons were incorporated into the PCR primers. Thus, for Pv200A, primers 5′-(dATGTCCCATGGAAACAGAAAGTTATAAGCAG)-3′ and 5′-(dCGCCCTCAACAAATCATAGTG)-3′ were used to amplify an *Nco*I/*Eco*RI fragment from the *Hin*dIII clone 4B-3-9. A pBluescript polylinker primer, 5′-(dGTGGATCCCCCGGGCTGCAGG)-3′ and the 3′-primer 5′-(dTTCCAAGGTCGACTATGGATTTTGCAAATCACCAAATGT)-3′ were used to amplify the contiguous *Eco*RI/*Sal*I fragment from the *Eco*RI clone 6.1-2. The PCR products were digested with the appropriate restriction enzymes and ligated into *Nco*I/*Sal*I digested pBS100 [7]. A *Bam*HI/*Sal*I fragment that contained the ADH2/GAPDH hybrid promoter product was exc [17] for yeast exp amplified as an A 6.1-2 using the ATGGCAGAAT and 5′-(dTTCC AGAAAACTCC cloning into pI above. pBS24 Pv200A and Pv2 into *Saccharomy* derivative of ABI *ADR1* gene [16]. gated and induce Yeast cell lysate SDS-PAGE gels and by immunob

Purification of cells that were harvested at 60 I in 1 M NaC Triton X-100 tion. Pv200A was pellet with 8 M The solution wa passed over a N exchange column HCl/1 mM EDT >95% pure by Coomasssie Blu Yeast cells expre and lysed in a sir extract was then a a Fast Flow Q column. The imr ing at 0.15–0.20 M was collected, co gel filtration on S in 4 M urea/10 m β-mercaptoethanc

Serology. Immu on monkey and essentially as desc obtained from monkeys that ha controls to assess

hybrid promoter [16] fused to the cloned PCR product was excised and cloned into pBS24 [17] for yeast expression. The Pv200B gene was amplified as an *NcoI*/*SalI* fragment from clone 6.1-2 using the primers 5'-(dCGGGAGCC-ATGGCAGAATCTGAGGCGCCTGAG)-3' and 5'-(dTTCCAAGGTCGACTAGCTAC-AGAAAACTCCCTCAAAGAG)-3'. Subcloning into pBS100 and pBS24 were as above. pBS24 plasmids containing the Pv200A and Pv200B genes were transformed into *Saccharomyces cerevisiae* strain JSC302, a derivative of AB116 [18] that overexpresses the *ADR1* gene [16]. Transformants were propagated and induced as described previously [7]. Yeast cell lysates were analyzed on 12.5% SDS-PAGE gels with Coomassie Blue staining, and by immunoblotting.

Purification of recombinant proteins. Yeast cells that were expressing Pv200A were harvested at 60 h and lysed with glass beads in 1 M NaCl/50 mM Tris/1 mM EDTA/0.1% Triton X-100 at pH 7.5 [7]. After centrifugation, Pv200A was extracted from the insoluble pellet with 8 M urea/100 mM Tris (pH 11.5). The solution was adjusted to pH 8.3 and passed over a Mono Q (Pharmacia) anion exchange column in 6 M urea/50 mM Tris-HCl/1 mM EDTA. Pv200A was found to be >95% pure by densitometric scanning of a Coomasssie Blue-stained SDS-PAGE gel. Yeast cells expressing Pv200B were harvested and lysed in a similar manner. The 8 M urea extract was then adjusted to pH 9 and run over a Fast Flow Q (Pharmacia) ion exchange column. The immunoblot-positive peak eluting at 0.15–0.20 M NaCl in a 0–1 M gradient was collected, concentrated and separated by gel filtration on Superose 12 resin (Pharmacia) in 4 M urea/10 mM Tris/1 mM EDTA/10 mM β-mercaptoethanol at pH 8.

Serology. Immunoblots and ELISA assays on monkey and human sera were performed essentially as described [13]. Monkey sera were obtained from *Saimiri boliviensis* squirrel monkeys that had been used previously as controls to assess the efficacy of a recombinant *P. vivax* circumsporozoite vaccine [8,9]. Anti-*Cynomolgus* conjugates were used to assay antibodies in monkey samples. Sera were taken 4 weeks after challenge with 10000 *P. vivax* Sal-1 sporozoites. Human sera were collected from a population in Brazil, outside the endemic zone, that had encountered a single 50-day exposure to *P. vivax* [19]. Serum samples were taken 8 months after the outbreak, which was completely controlled by chemotherapy and insecticides. Enhancement of sensitivity in the human sera ELISAs was achieved using the 3, 3', 5, 5'-tetramethylbenzidine (TMB) as substrate and reagents as recommended by the manufacturer (Kirkegaard and Perry Laboratories Inc., Gaithersburg, MD).

Results and Discussion

Molecular cloning of the Pv200 gene. Overlapping λZAPII phage clones were isolated that contained the large open reading frame encoding Pv200 (Fig. 1). Most of the Pv200 protein sequence is encoded by the *Eco*RI clone (6.1-2). The amino-terminal region is contained within the ca. 7-kb *Hin*dIII clone (4B-3-9). The full DNA sequence of a 5.83-kb composite *Nde*I fragment that includes the entire Pv200 coding sequence is shown (Fig. 2). Typical promoter elements [20] are found within the 5'-region of the sequence, and a consensus motif for efficient initiation of translation [21] is apparent around the proposed initiation codon.

Structure of Pv200. As deduced from the composite sequence, Pv200 is derived from a 1751-amino acid precursor protein that contains a typical amino-terminal secretory signal sequence [22], and a transmembrane domain (TMD) (Fig. 1). A predicted signal peptidase cleavage site [22] occurs after Cys19. Also, within the hydrophobic TMD (Fig. 2), is a stretch of 4 serine residues that is reminiscent of the signal for attachment of glycosylphosphatidyl inositol anchor sequences to proteins of various protozoans, including Pf195 [23].

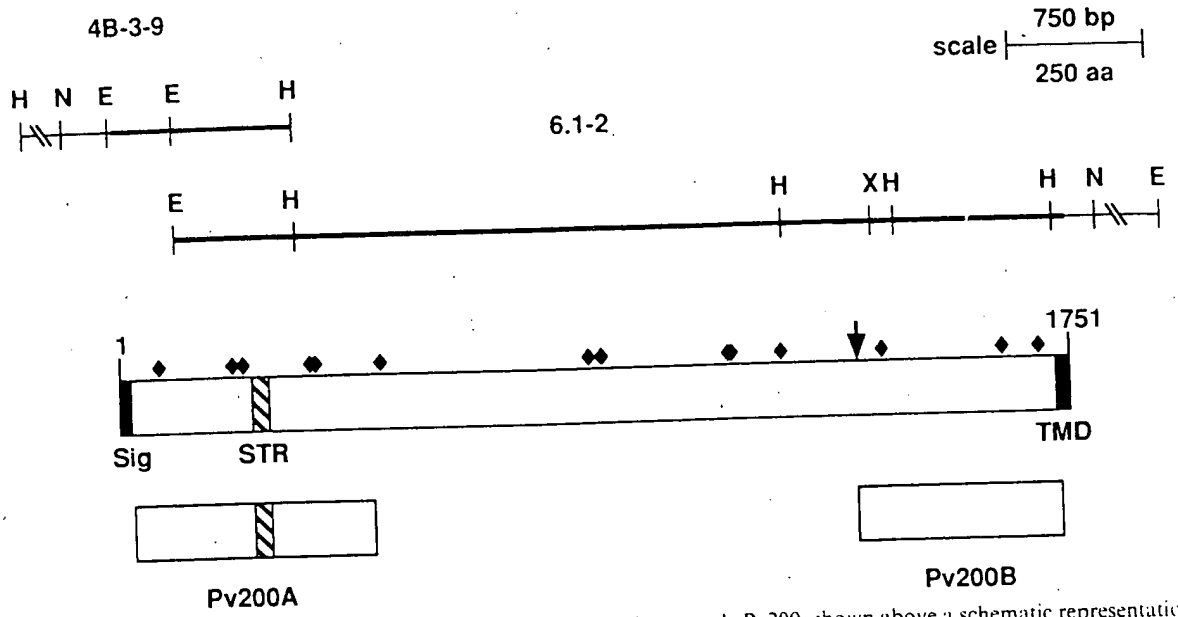
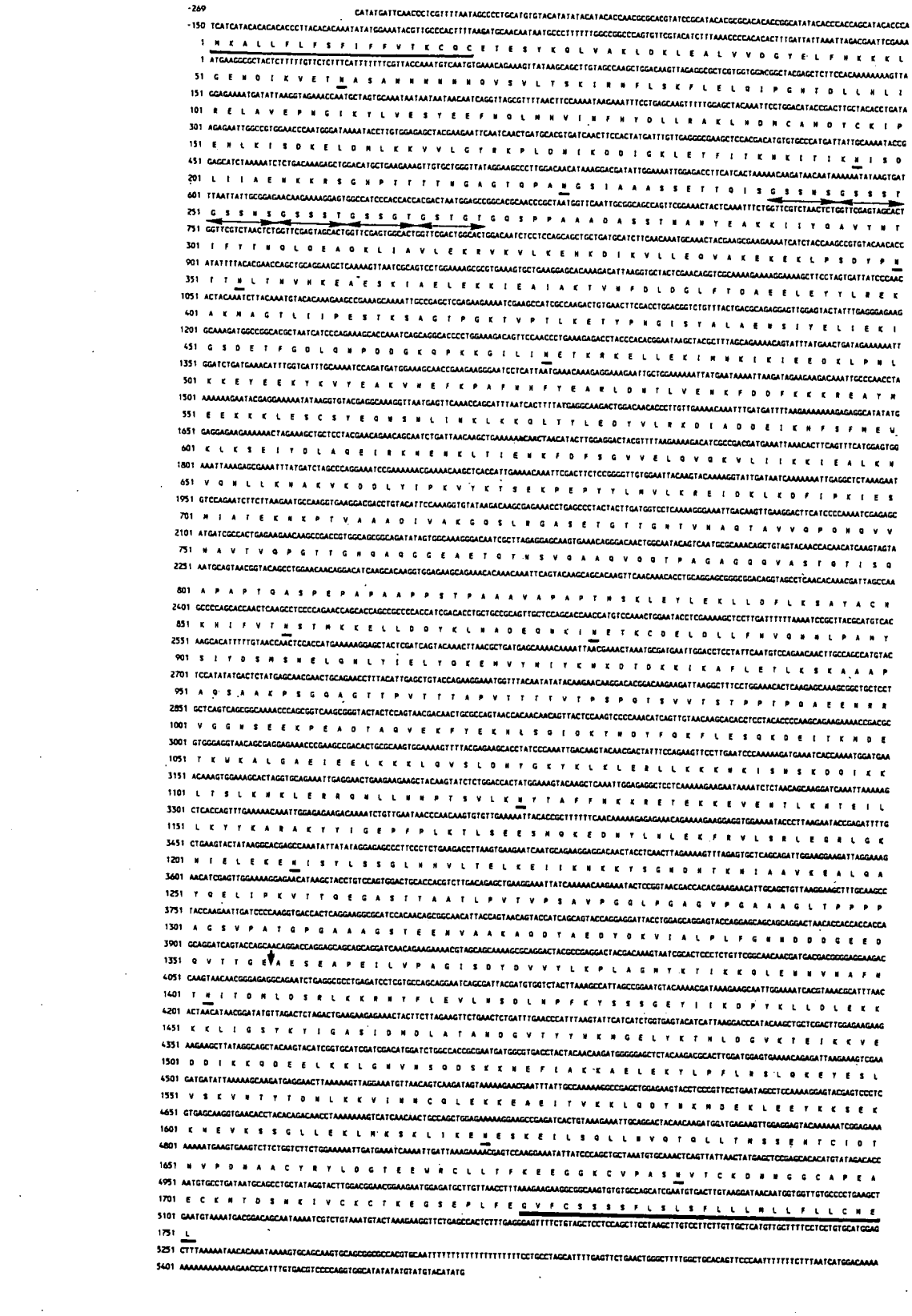


Fig. 1. Restriction maps of the overlapping genomic DNA clones that encode Pv200, shown above a schematic representation of the encoded protein sequence. The flanking *Nde*I sites (Fig. 2, legend) are shown (N). Other sites used for mapping and DNA sequencing: *Hin*dIII (H), *Eco*RI (E), and *Xba*I (X). Potential asparagine-linked glycosylation sites are shown (◆), as is the serine/threonine-rich region (STR). The signal sequence (Sig) and putative transmembrane domain (TMD) are shown in black, and a potential proteolytic cleavage site (see text) is arrowed. Also shown schematically, are the regions Pv200A and Pv200B that were selected for expression in yeast.

Overall amino acid sequence identity between our Pv200 and Pf195 from various sources [24] is in the 36–37% range, although a region of more extensive identity (45%) exists in the amino-terminal region [10, and present study]. Similarly, amino acid sequence identities between Pv200 and the major merozoite surface antigens of *P. chabaudi chabaudi* and *P. yoelii yoelii* were 34.6% and 34.9%, respectively [25,26]. Also present is a potential proteolytic cleavage site identical to that defined previously for the generation of the carboxy-terminal p42 protein of the Pf195 precursor [27] (Figs. 1 and 2, arrowed). Also in common with Pf195, Pv200 does not contain multiple repetitive elements that are characteristic of many other proteins of malaria parasites. Six serine/threonine-rich motifs of 5 amino acids are noted, however (Figs. 1, 2). Also, the striking 23-glutamine residue repeat of the Belèm strain Pv200 protein was not present in the sequence encoded by our cloned gene, but rather, was replaced by a 35-amino acid residue stretch that contained only 6 glutamine residues. This, together with several other smaller insertions and deletions accounts for the larger size of the Sal-1 Pv200 precursor over that of the 1726 amino acid Pv200 precursor from the Belèm strain. Despite this size difference, the two Pv200 precursors are relatively well conserved, with an overall amino acid sequence identity of 81%. As with the major merozoite surface antigens from all *Plasmodium* species that have been studied thus far, this intra-species homology could be divided into areas of the protein of relatively

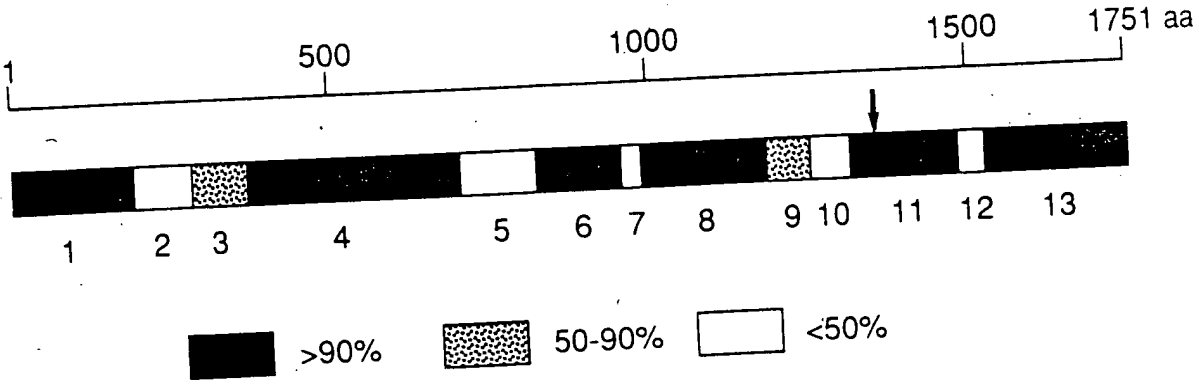
Fig. 2. Composite DNA sequence of a 5.83-kb *Nde*I fragment containing the Pv200 coding sequence. The 1751-amino acid large open reading frame is shown with proposed amino-terminal signal and carboxy-terminal membrane-spanning sequences underlined. Also underlined, are 13 asparagine residues that are potential sites for glycosylation. A 30-amino acid serine/threonine-rich region between amino acids 241 and 270 that contains 6 copies of the 5 amino acid repeat motif G.S.(S/T).(N S G).(S/T) is also noted (double arrow). A potential proteolytic cleavage site (see text) after amino acid Glu1356 is arrowed.



low amino acid sequence identity separated by conserved or semi-conserved regions [11,24]. We have identified thirteen distinct regions of Pv200 that are typified by their amino acid sequence identity levels between the Sal-1 and Belèm strains (Fig. 3). Although not fully comparable with the Pf195 variable blocks, or their junctions [11,24], the existence of this distinct pattern suggests the possibility that Pv200 diversity could be generated by intragenic recombination of a limited number of alleles, as is the case with Pf195 [24,28]. Of further note, the carboxy-terminal region of the Sal-1 strain Pv200, equivalent to p42 of *P. falciparum*, is encoded to a large degree by 2 of the highly conserved gene segments (11 and 13) and exhibits 91% amino acid sequence identity with the corresponding region of the Belèm strain protein.

Expression of Pv200 in yeast. Two domains of Pv200 were selected for expression studies (Fig. 1). The first, designated Pv200A, includes amino acids 20–462 of the Pv200 precursor. Pv200A represents, therefore, a protein of approx. 49 kDa from the amino-terminal region of Pv200. A similarly expressed region of Pf195 can elicit good immunological responses against native Pf195 in mice and rabbits (these authors, S.P. Chang, G.S. Hui, unpublished observations). Furthermore, this region of Pf195, expressed in bacteria, has been shown previously to induce partial protection in *Aotus* monkeys that were subjected to blood stage challenge with *P. falciparum* [5]. The second domain that was expressed, Pv200B, containing amino acids 1357–1729 of the Pv200 precursor, is the homolog of the carboxy-terminal fragment of Pf195 that has been structurally defined as p42 [27]. This protein is of considerable importance since the carboxy-terminus of the major merozoite surface antigen has been implicated in the induction of a protective immune response against *P. falciparum* infection [26,29]. Additional studies on this *P. falciparum* antigen have indicated that secreted recombinant p42, from insect cells, is recognized by conformation-dependent antibodies [3].

Each gene construct was generated by PCR [15] and expressed in the yeast *S. cerevisiae*. Pv200A was produced at particularly high



REGION	IDENTITY	REGION	IDENTITY	REGION	IDENTITY
1	92%	6	92%	11	96%
2	46%	7	17%	12	37%
3	67%	8	95%	13	99%
4	97%	9	63%		
5	39%	10	48%		

Fig. 3. Schematic representation of amino acid sequence identity distributions between Sal-1 and Belèm strain [11] Pv200 proteins. Shown here is the larger Sal-1 strain protein, with gaps scored as mismatches. The position of the putative proteolytic cleavage site for generation of the *P. vivax* p42 homologue is shown (arrowed).

66 –

43 –

31 –

1

Fig. 4. SDS-PAGI
plasmids were lyse
0.2% bromopheno
proteins visualized
yeast cell lysates; l
purified Pv200A a
presta

levels (Fig. 4A,
the insoluble fr
detergent-conta
be purified to >
scheme. Pv200
levels (<20%
buffer insoluble
to greater than
procedure. The

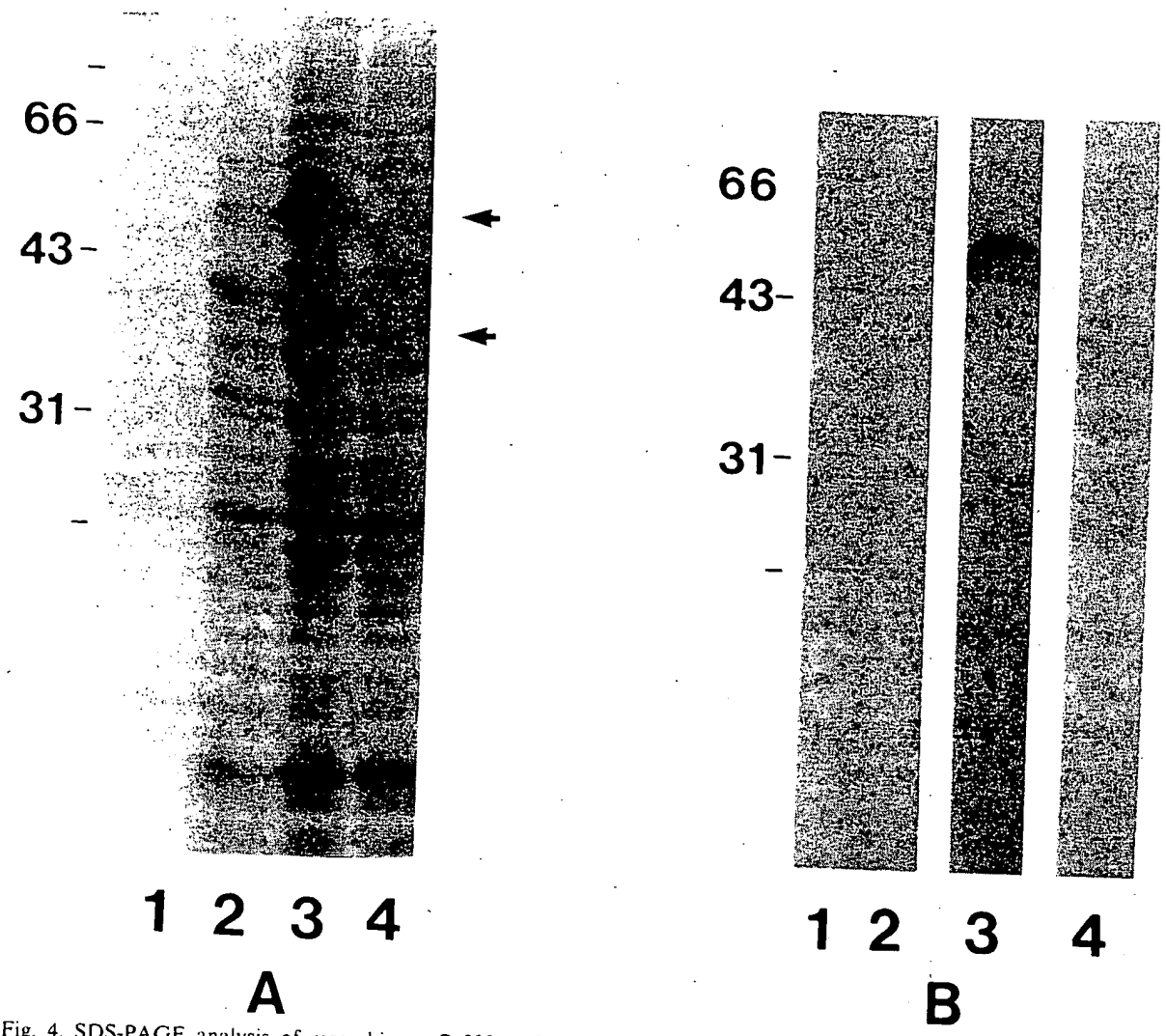


Fig. 4. SDS-PAGE analysis of recombinant Pv200 antigens. (A) Yeast cells transformed with control or recombinant plasmids were lysed in gel loading buffer containing 10% glycerol, 50 mM DTT, 3% SDS, 625 mM Tris-HCl pH 6.8, and 0.2% bromophenol blue. After 3 sequential boiling and freezing steps, the lysates were electrophoresed on 12.5% gels and proteins visualized by staining with Coomassie Blue. Lane 1, molecular weight markers (BioRad prestained); lane 2, control yeast cell lysates; lanes 3, 4, total lysates from yeast cells expressing Pv200A and Pv200B respectively. (B) Immunoblot of purified Pv200A and Pv200B using pooled human sera (see text for details). Lane 1, molecular weight markers (BioRad prestained); lane 2, control yeast cell lysate; lanes 3,4; purified Pv200A and Pv200B, respectively.

levels (Fig. 4A, lane 3) and represented 81% of the insoluble fraction, after lysis in a nonionic detergent-containing buffer. This protein could be purified to >95% by a one step purification scheme. Pv200B was expressed at moderate levels (<20% of total yeast Triton X-100 buffer insoluble fraction), but could be purified to greater than 85% using a 2-step purification procedure. The purified Pv200A and Pv200B proteins were shown to recognize antibodies in sera from individuals with a previous history of *P. vivax* infection [19]. Pooled human sera from such individuals, who were positive in their responses to the *P. vivax* CS protein (see below), was shown to further react with the recombinant merozoite surface antigens, by immunoblot analysis (Fig. 4B).

Seroreactivity of recombinant Pv200 antigens. Pv200A and Pv200B were used to analyze sera from *Saimiri* monkeys that were infected previously with the Sal-1 strain of *P. vivax* by sporozoite challenge [8,9]. Somewhat surprisingly, the ELISA titers from these animals were low and were not indicative of disease state. In only 2 animals did we observe high titer responses, and in each case this was against Pv200B (data not shown). The reasons for this absence of high titer responses are currently unknown. However, these relatively low overall titers may reflect the fact that this was the primary challenge of a group of monkeys that had not been exposed previously to *P. vivax* malaria. In contrast, we observed relatively high ELISA titers in a human population that had been subject to a single outbreak of vivax malaria (Fig. 5). ELISA titers against both Pv200A and Pv200B were above control values, as were titers against the recombinant CS protein (Vivax2) [30]. Noticeably, titers against Pv200B were consistently higher than those against Pv200A. The most direct explanation for this is that the carboxy-terminal region of Pv200 is simply more immunogenic than the amino-terminal domain, at least in a primary infection in humans. Alternatively, Pv200B might possess greater conformational integrity than Pv200A when compared with the corresponding regions of the native proteins, and might thus be more antigenic in the ELISA format. A third explanation is that higher titers against Pv200B could reflect the more conserved structure of this region. For example, the region of the Sal-1 Pv200 protein defined by our Pv200A molecule shares 76% amino acid sequence identity with the corresponding region of the Belém strain Pv200 protein whereas, as mentioned above, the carboxy-terminal regions, corresponding to Pv200B, are 91% identical. In conclusion, we have demonstrated that recombinant Pv200 proteins produced in recombinant yeast are able to recognize antibodies in infected monkeys and humans. The production of these proteins in large quantities will allow further and more detailed studies of their antigenicity, and their potential use as diagnostic reagents. In addition, their ability to elicit protective immune responses in experimental animals can now be evaluated.

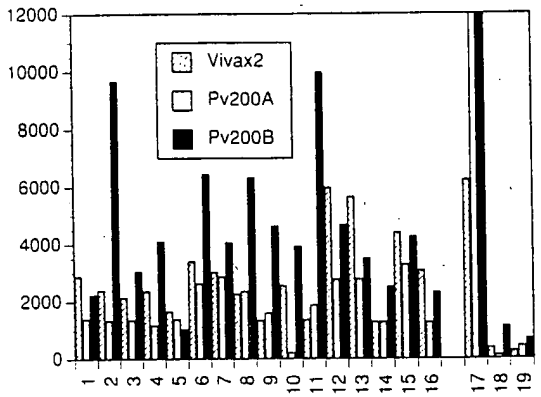


Fig. 5. ELISA titers of human sera. Sera from individuals with overt infections, in which parasitemia was detected by thick blood smears, and who were treated with oral chloroquine (numbers 1–16). Three individuals were also included as controls, one positive (number 17) who had had multiple *P. vivax* and *P. falciparum* infections, with titers of around 655000 and 24000 for Pv200A and B respectively, and 2 negative (numbers 18 and 19) who were never exposed to *P. vivax* infections. Individuals in each group were assayed for reactivity against the *P. vivax* CS protein using a recombinant CS protein (Vivax2) [30] as well as against Pv200 A and B.

Acknowledgements

We thank our friends and collaborators for their encouragement and support. In particular, we thank Drs. Pablo Valenzuela and Lou Miller. We also thank Dr. Peter David for communication of results prior to publication, and Peter Anderson for preparation of the manuscript. This work was supported by Chiron Corporation and Ciba Geigy Ltd.

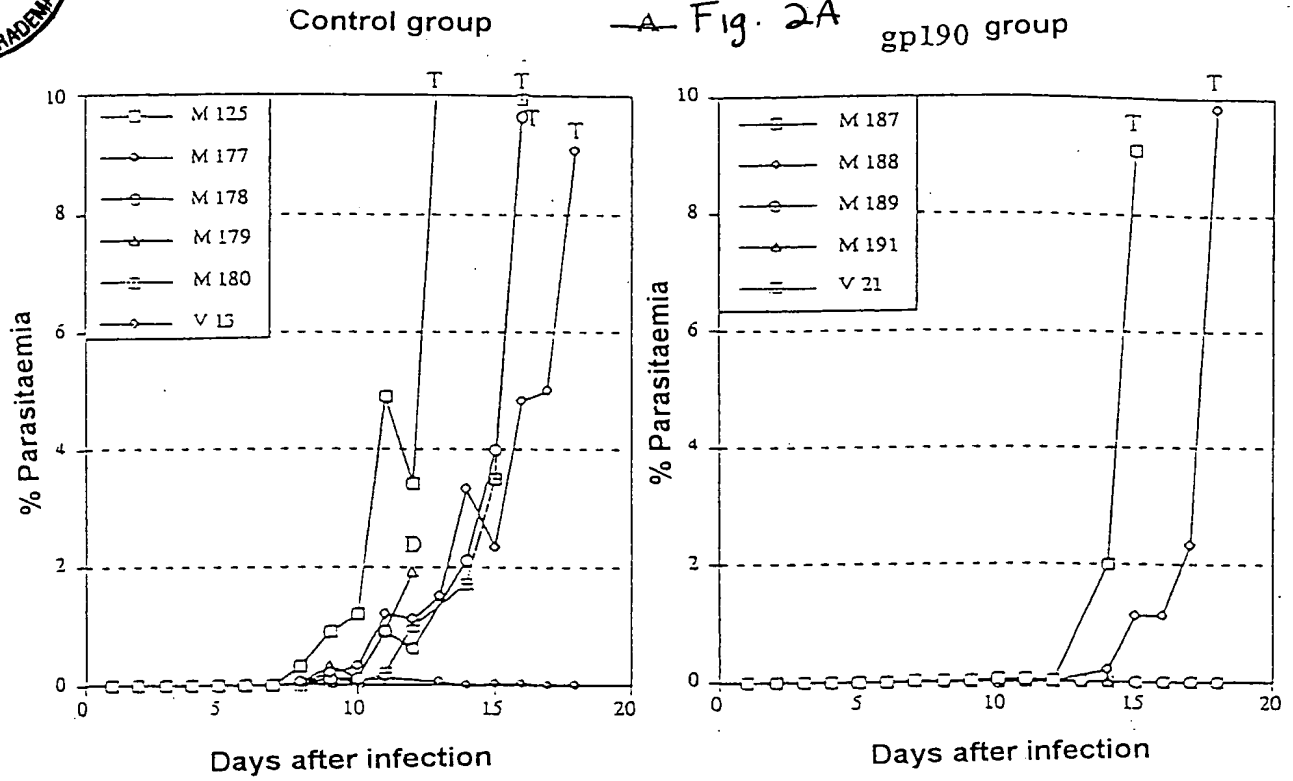
References

1 Holder, A.A. and Freeman, R.R. (1982) Biosynthesis and processing of a *Plasmodium falciparum* schizont antigen recognized by immune serum and a monoclonal antibody. J. Exp. Med. 156, 1528–1538.
2 Siddiqui, W.A., Tam, L.Q., Kramer, K.J., Hui, G.S.N., Case, S.E., Yamaga, K.M., Chang, S.P., Chan, E.B.T. and Kan, S.-C. (1987) Merozoite surface coat precursor protein completely protects *Aotus* monkeys against *Plasmodium falciparum* malaria. Proc. Natl. Acad. Sci. USA 84, 3014–3018.
3 Murphy, V.F., A.A. (1990) Ex insect cells and secretion. Paras
4 Müller, H.-M., Lombardi, S., Development of the major surf *falciparum*. Infec
5 Herrera, S., He Caspers, P., Dö (1990) Immuniz *dium falciparum* Proc. Natl. Acad
6 Arnot, D.E., Ba V., Nussenzweig, zoite protein of characterization Science 230, 815-
7 Barr, P.J., Gib Hollingdale, M.F sion in yeast of a use in a human 1160–1171.
8 Collins, W.E., N bush, T.K., Nard Young, J.F., Wa H.L., Barr, P.J. Broderson, J.R., V.K. and Wilson, *sciureus boliviensi* based on the circ *vivax*. Am. J. Tro
9 Collins, W.E., N Bathurst, I.C., Ni G.H., Barr, P.J., E V.K., Stanfill, P. (1990) Further st *sciureus boliviensis* the circumsporoz J. Trop. Med. Hy
10 del Portillo, H.A., Udagama, P.V., N *Plasmodium vivax* blood-stage surfac 353.
11 del Portillo, H.A., P.H. (1991) Prima antigen 1 of *Pl* conserved between Natl. Acad. Sci. U
12 Feinberg, A. and V radiolabeling DNA to high specific act
13 Sambrook, J., Ma Molecular Cloning Cold Spring Harb Harbor, NY.
14 Sanger, F., Nicklen sequencing with c Natl. Acad. Sci. U
15 Saiki, R.K., Scharf, G.T., Erlich, H.A. amplification of restriction site an anemia. Science 23

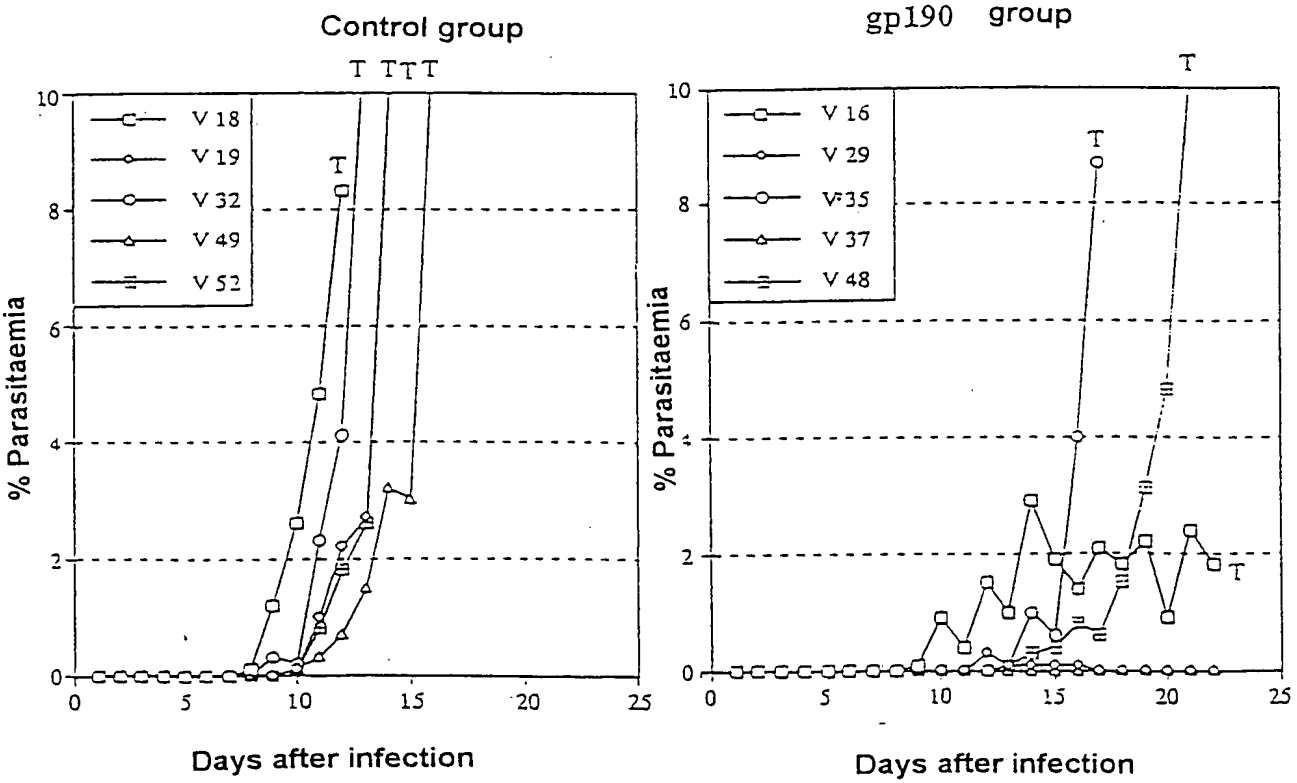
3 Murphy, V.F., Rowan, W.C., Page, M.J. and Holder, A.A. (1990) Expression of hybrid malaria antigens in insect cells and their engineering for correct folding and secretion. Parasitology 100, 177–183.
4 Müller, H.-M., Früh, K., von Brunn, A., Esposito, F., Lombardi, S., Crisanti, A. and Bujard, H. (1989) Development of the human immune response against the major surface protein (gp190) of *Plasmodium falciparum*. Infect. Immun. 57, 3765–3769.
5 Herrera, S., Herrera, M.A., Perlaza, B.L., Burki, Y., Caspers, P., Döbeli, H., Rotmann, D. and Certa, U. (1990) Immunization of *Aotus* monkeys with *Plasmodium falciparum* blood-stage recombinant proteins. Proc. Natl. Acad. Sci. USA 87, 4017–4021.
6 Arnot, D.E., Barnwell, J.W., Tam, J.P., Nussenzweig, V., Nussenzweig, R.S. and Enea, V. (1985) Circumsporozoite protein of *Plasmodium vivax*: gene cloning and characterization of the immunodominant epitope. Science 230, 815–818.
7 Barr, P.J., Gibson, H.L., Enea, V., Arnot, D.E., Hollingdale, M.R. and Nussenzweig, V. (1987) Expression in yeast of a *Plasmodium vivax* antigen of potential use in a human malaria vaccine. J. Exp. Med. 165, 1160–1171.
8 Collins, W.E., Nussenzweig, R.S., Ballou, W.R., Ruebush, T.K., Nardin, E.H., Chulay, J.D., Majarian, J.F., Young, J.F., Wasserman, G.F., Bathurst, I.C., Gibson, H.L., Barr, P.J., Hoffman, S.L., Wasserman, S.S., Broderson, J.R., Skinner, J.C., Procell, P.M., Filipski, V.K. and Wilson, C.L. (1989) Immunization of *Saimiri sciureus boliviensis* monkeys with recombinant vaccines based on the circumsporozoite protein of *Plasmodium vivax*. Am. J. Trop. Med. Hyg. 40, 455–464.
9 Collins, W.E., Nussenzweig, R.S., Ruebush, T.K., Bathurst, I.C., Nardin, E.H., Gibson, H.L., Campbell, G.H., Barr, P.J., Broderson, J.R., Skinner, J.C., Filipski, V.K., Stanfill, P.S., Roberts, J.M. and Wilson, C.L. (1990) Further studies on the immunization of *Saimiri sciureus boliviensis* with recombinant vaccines based on the circumsporozoite protein of *Plasmodium vivax*. Am. J. Trop. Med. Hyg. 43, 576–583.
10 del Portillo, H.A., Gysin, J., Mattei, D.M., Khouri, E., Udagama, P.V., Mendis, K.N. and David, P.H. (1988) *Plasmodium vivax*: cloning and expression of a major blood-stage surface antigen. Exp. Parasitol. 67, 346–353.
11 del Portillo, H.A., Longacre, S., Khouri, E. and David, P.H. (1991) Primary structure of the merozoite surface antigen 1 of *Plasmodium vivax* reveals sequences conserved between different *Plasmodium* species. Proc. Natl. Acad. Sci. USA 88, 4030–4034.
12 Feinberg, A. and Volgelstein, B. (1984) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. 137, 266–267.
13 Sambrook, J., Maniatis, T. and Fritsch, E.F. (1989) Molecular Cloning. A Laboratory Manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
14 Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74, 5463–5467.
15 Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A. and Arnheim, N. (1985) Enzymatic amplification of β globin genome sequences and restriction site analysis for diagnosis of sickle cell anemia. Science 230, 1350–1354.
16 Shuster, J.R. (1989) Regulated transcriptional systems for the production of proteins in yeast: regulation by carbon source. In: Yeast Genetic Engineering (Barr, P.J., Brake, A.J. and Valenzuela, P., eds.), pp. 83–108, Butterworth, Stoneham, MA.
17 Sabin, E.A., Lee-Ng, C.T., Shuster, J.R. and Barr, P.J. (1989) High-level expression and in vivo processing of chimeric ubiquitin fusion proteins in *Saccharomyces cerevisiae*. Bio/Technology 7, 705–709.
18 Barr, P.J., Cousens, L.S., Lee-Ng, C.T., Medina-Selby, A., Masiarz, F.R., Hallewell, R.A., Chamberlain, S.H., Bradley, J.D., Lee, D., Steimer, K.S., Poulter, L., Burlingame, A.L., Esch, F. and Baird, A. (1988) Expression and processing of biologically active fibroblast growth factors in the yeast *Saccharomyces cerevisiae*. J. Biol. Chem. 263, 16471–16478.
19 Fontes, C.J., Bathurst, I. and Krettli, A.U. (1991) *Plasmodium vivax* sporozoite antibodies in individuals exposed during a single malaria outbreak in a nonendemic area. Am. J. Trop. Med. Hyg. 44, 28–33.
20 Lewis, A.P. (1990) Sequence analysis upstream of the gene encoding the precursor to the major merozoite surface antigens of *Plasmodium yoelii*. Mol. Biochem. Parasitol. 39, 285–288.
21 Kozak, M. (1984) Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. Nucleic Acids Res. 12, 857–872.
22 von Heijne, G. (1986) A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. 14, 4683–4690.
23 Smythe, J.A., Coppel, R.L., Brown, G.V., Ramasamy, R., Kemp, D.J. and Anders, R.F. (1988) Identification of two integral membrane proteins of *Plasmodium falciparum*. Proc. Natl. Acad. Sci. USA 85, 5195–5199.
24 Tanabe, K., MacKay, M., Goman, M. and Scaife, J.G. (1987) Allelic dimorphism in a surface antigen gene of the malaria parasite *Plasmodium falciparum*. J. Mol. Biol. 195, 273–287.
25 Deleersnijder, W., Hendrix, D., Bendahman, N., Hanegreefs, J., Brijs, L., Hamers-Casterman, C. and Hamers, R. (1990) Molecular cloning and sequence analysis of the gene encoding the major merozoite surface antigen of *Plasmodium chabaudi chabaudi* IP-PC1. Mol. Biochem. Parasitol. 43, 231–244.
26 Holder, A.A. and Freeman, R.R. (1984) Characterisation of a high molecular weight protective antigen of *Plasmodium yoelii*. Parasitol. 88, 211–219.
27 Heidrich, H.-G. (1988) Isolation and functional characterization of *Plasmodium falciparum* merozoite antigens. Biol. Cell 64, 205–214.
28 Peterson, M.G., Coppel, R.L., Moloney, M.B. and Kemp, D.J. (1988) Third form of the precursor to the major merozoite surface antigens of *Plasmodium falciparum*. Mol. Cell. Biol. 8, 2664–2667.
29 Burns, J.M., Jr., Daly, T.M., Vaidya, A.B. and Long, C.A. (1988) The 3′ portion of the gene for a *Plasmodium yoelii* merozoite surface antigen encodes the epitope recognized by a protective monoclonal antibody. Proc. Natl. Acad. Sci. USA 85, 602–606.
30 Bathurst, I.C., Gibson, H.L., Kansopon, J., Hahm, B.K., Hollingdale, M.C. and Barr, P.J. (In Press) Muramyl peptide adjuvants for *Plasmodium falciparum* and *P. vivax* circumsporozoite vaccines in rodent model systems. Biotech. Ther.

~~Fig. 2~~

~~A~~ Fig. 2A

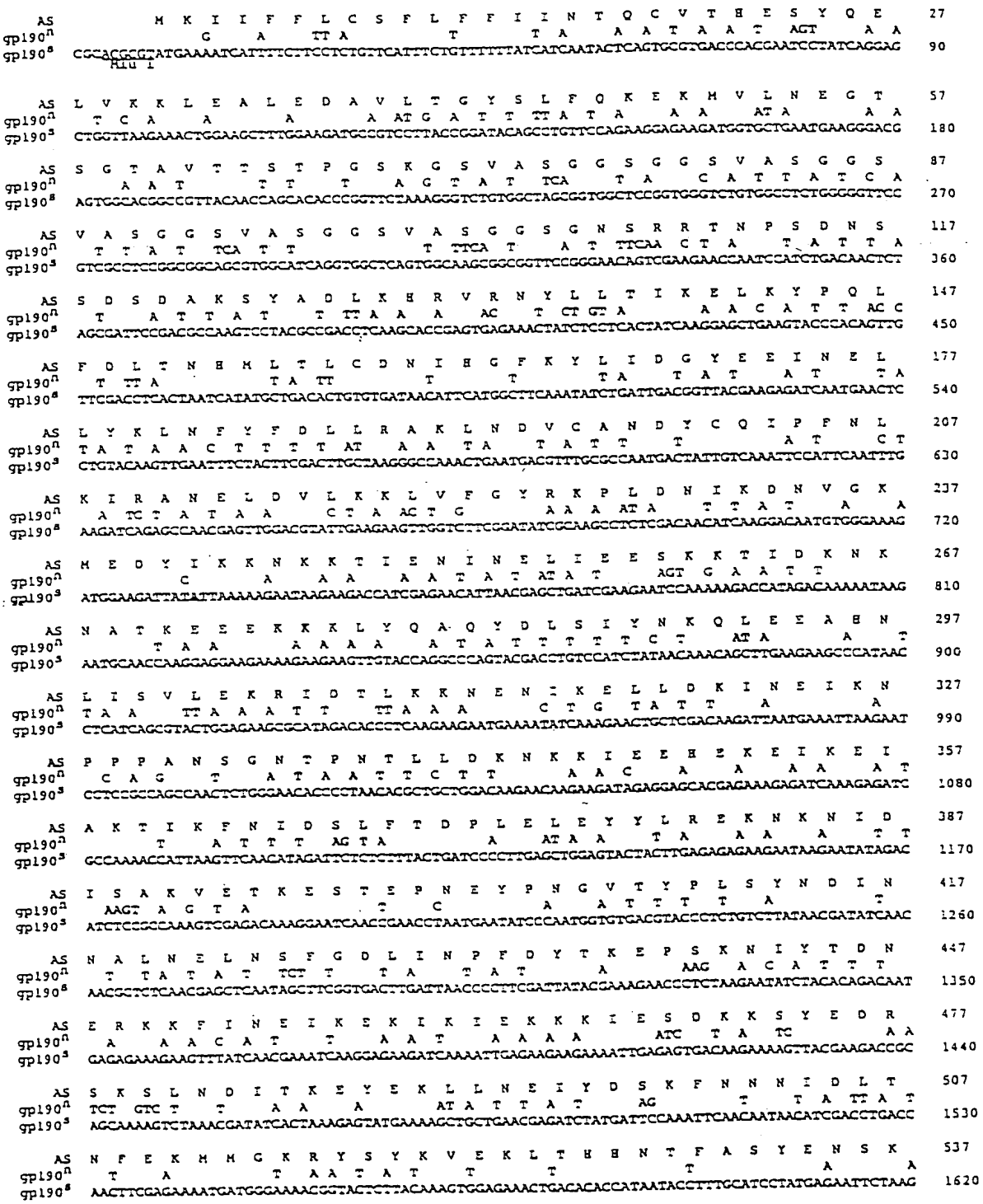


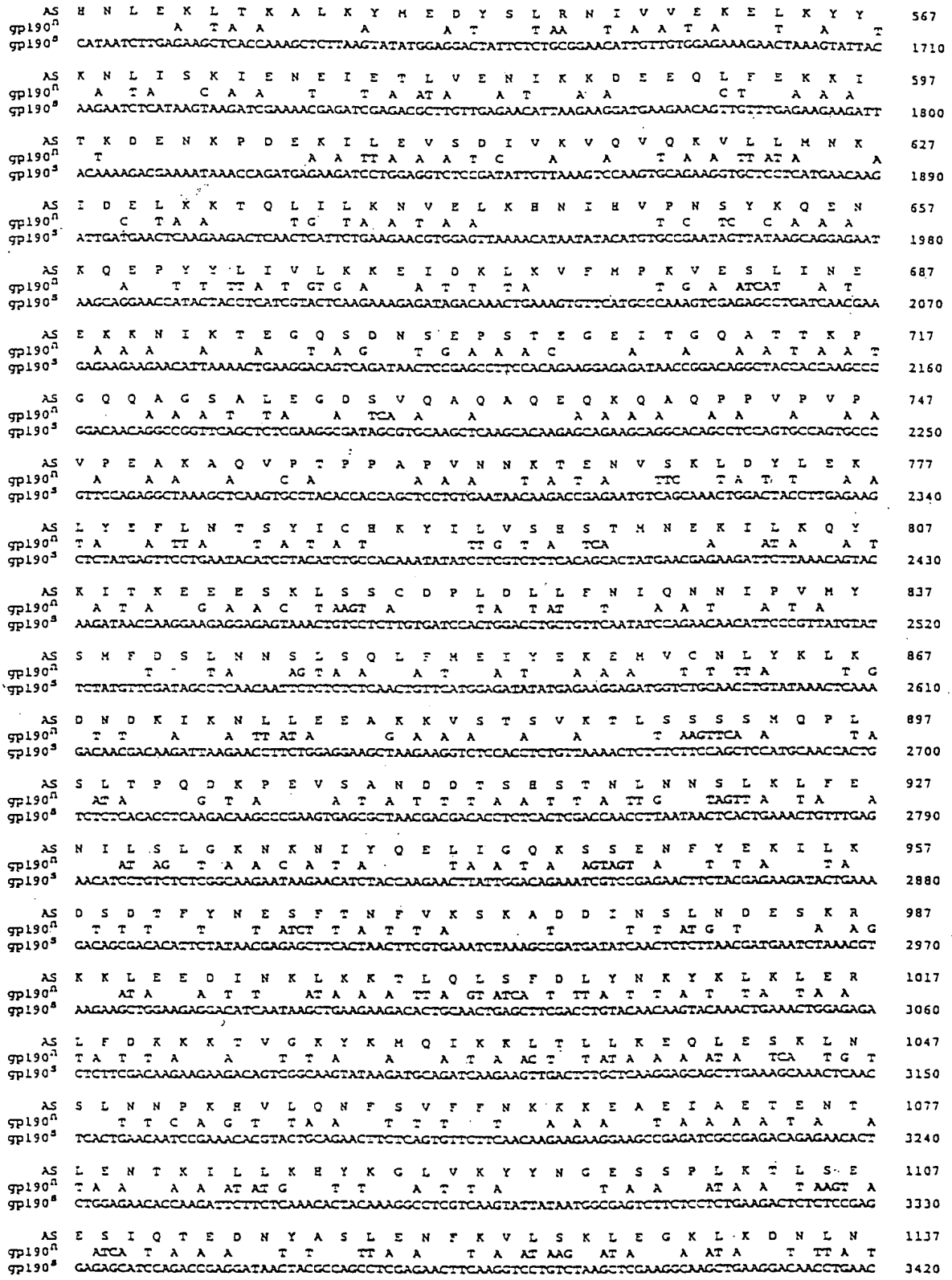
~~B~~ Fig. 2B

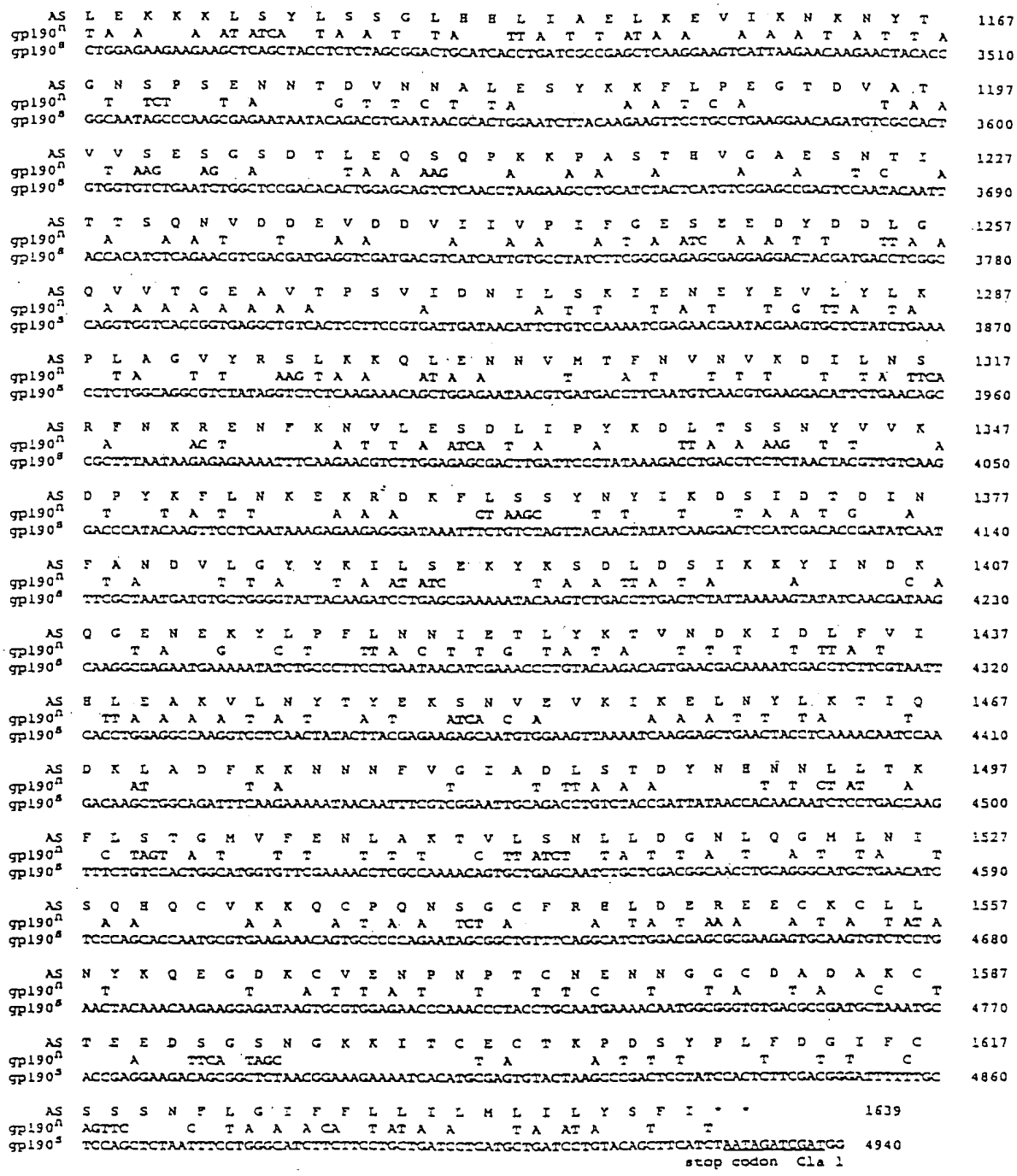


~~Fig. 3e~~ Fig. 3c - Fig. 3x 5/16

DNA sequence of the native (gp190n) and of the synthetic gene (gp190s) for gp190 from FCG-1







~~Fig. 3D~~
Fig. 3Y

N′-terminus C′-terminus

gp190^{S1} Sequence

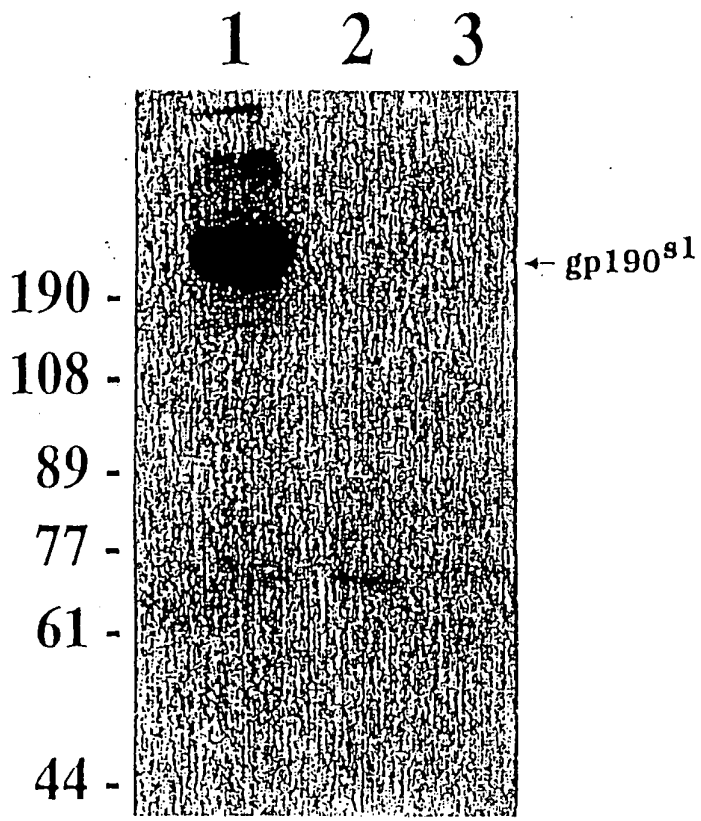
DNA Sequence	GC ACGCGTATGAAAATC	---------- AGCTCTAATTAATAGGCGGCCGCATCGATGGC
AA Sequence	Mlu I Met Lys Ile	---------- Ser Ser Asn stop codon Not I Cla I
AA Position	1 2 3	---------- 1619 1620 1621

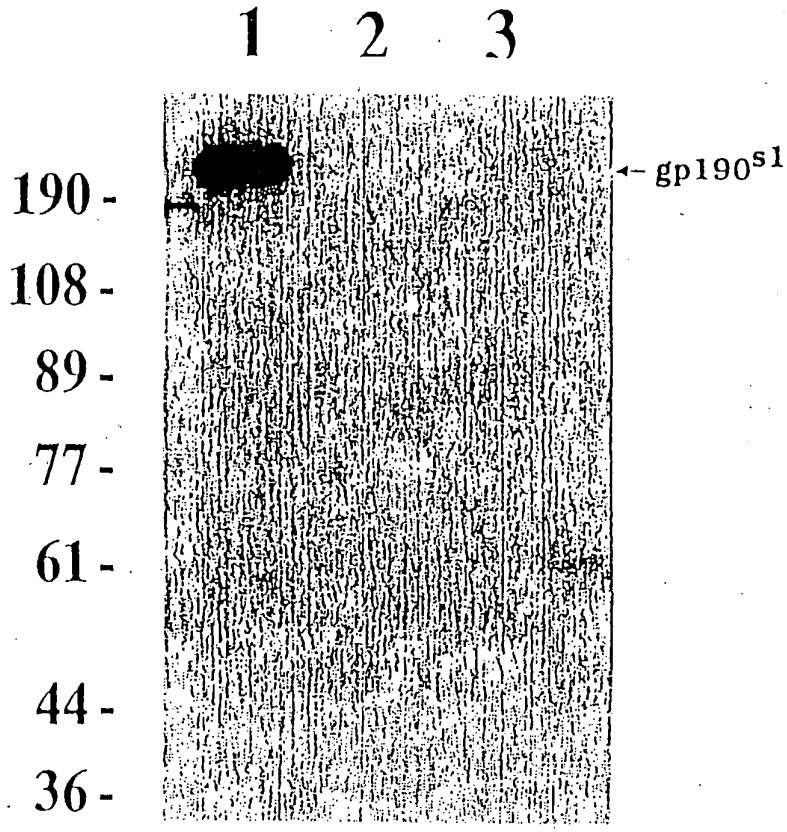
gp190^{S2} Sequence

DNA Sequence	GC GGATCCGTGACCCAC	---------- AGCTCTAATTAATAGGCGGCCGCATCGATGGC
AA Sequence	BamH I Val Thr His	---------- Ser Ser Asn stop codon Not I Cla I
AA Position	20 21 22	---------- 1619 1620 1621

8/16

~~Fig. 5C~~
Fig. 5B





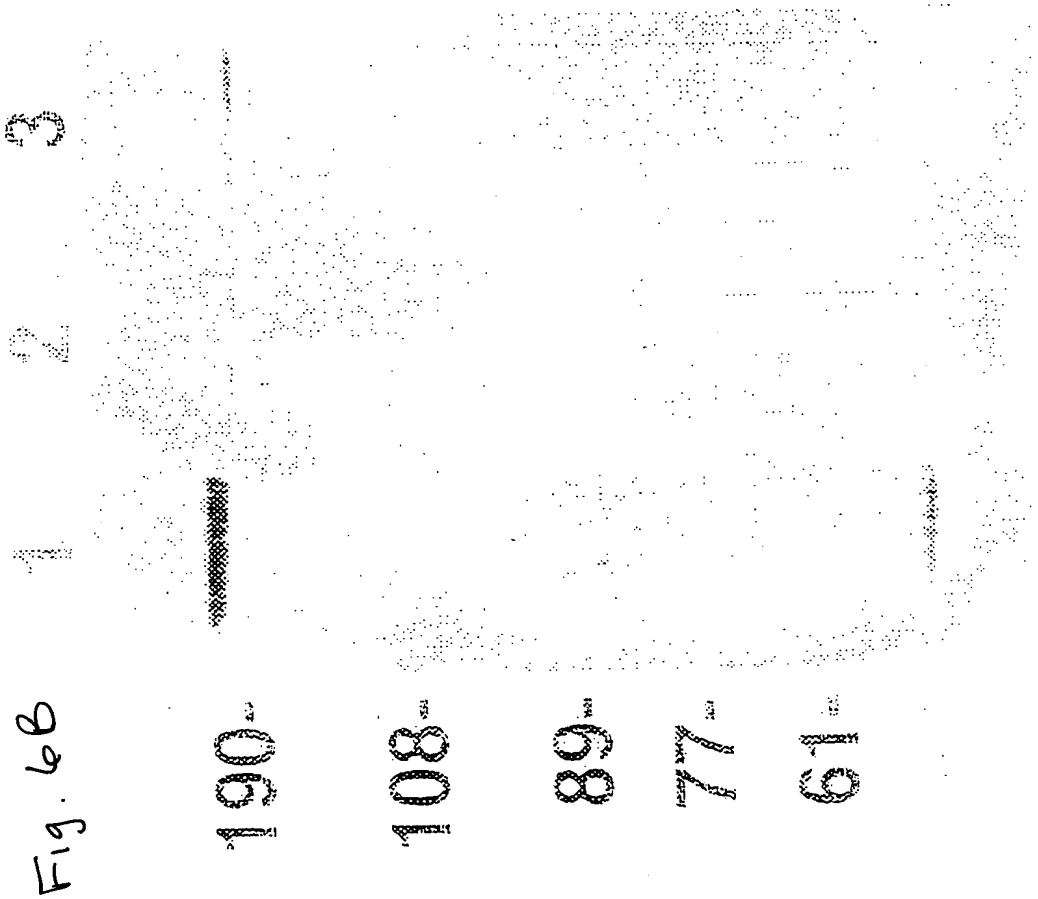


Fig. 6B

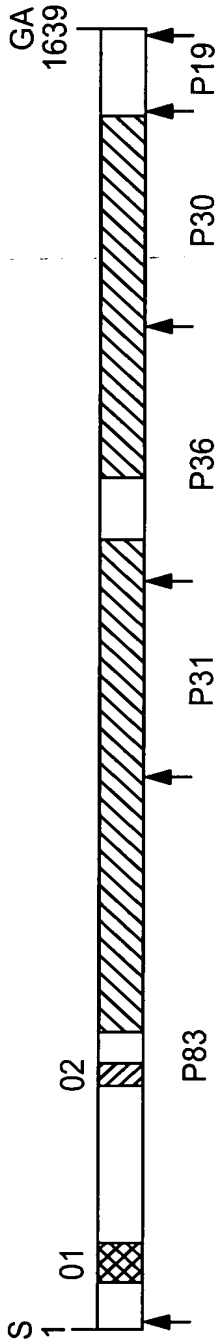
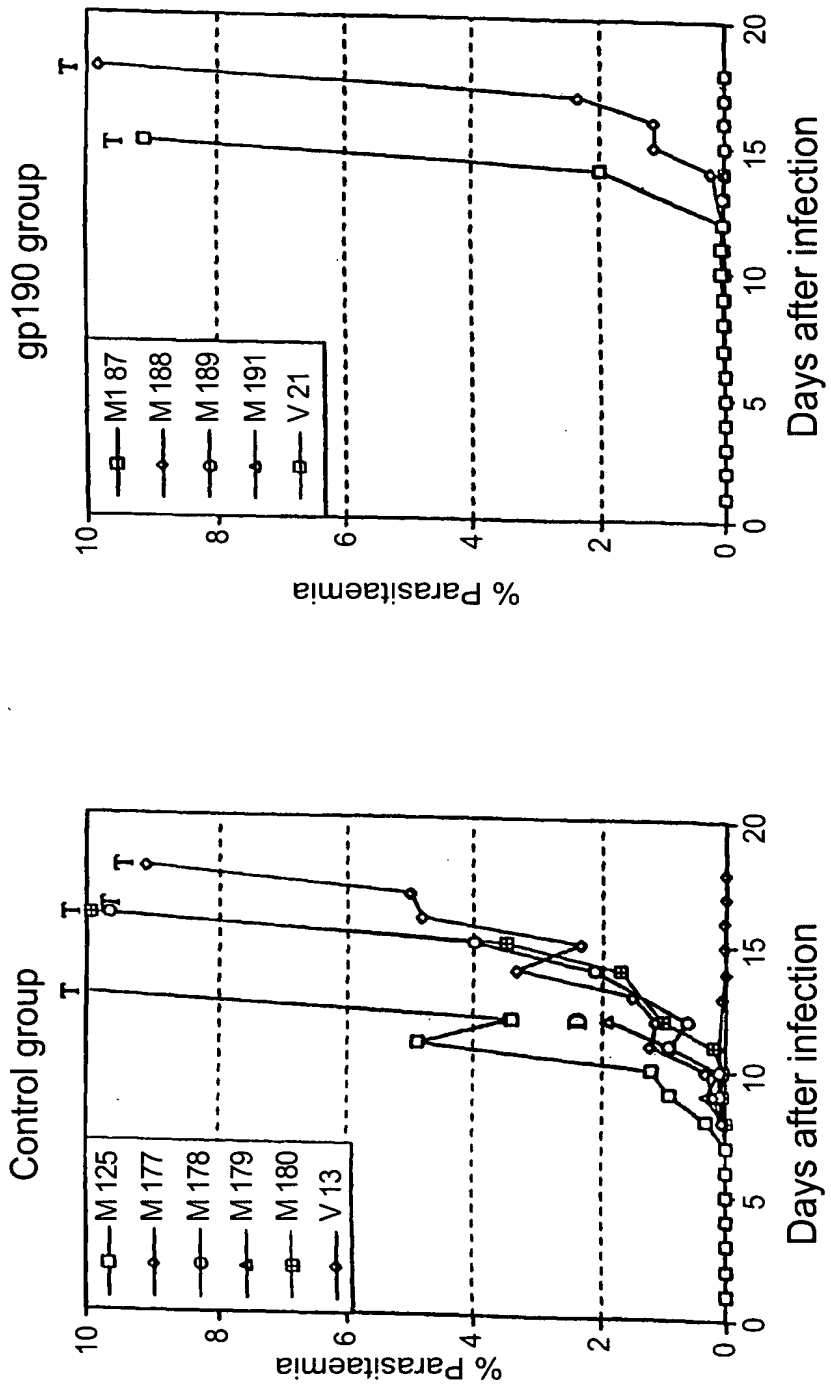


FIG. 1

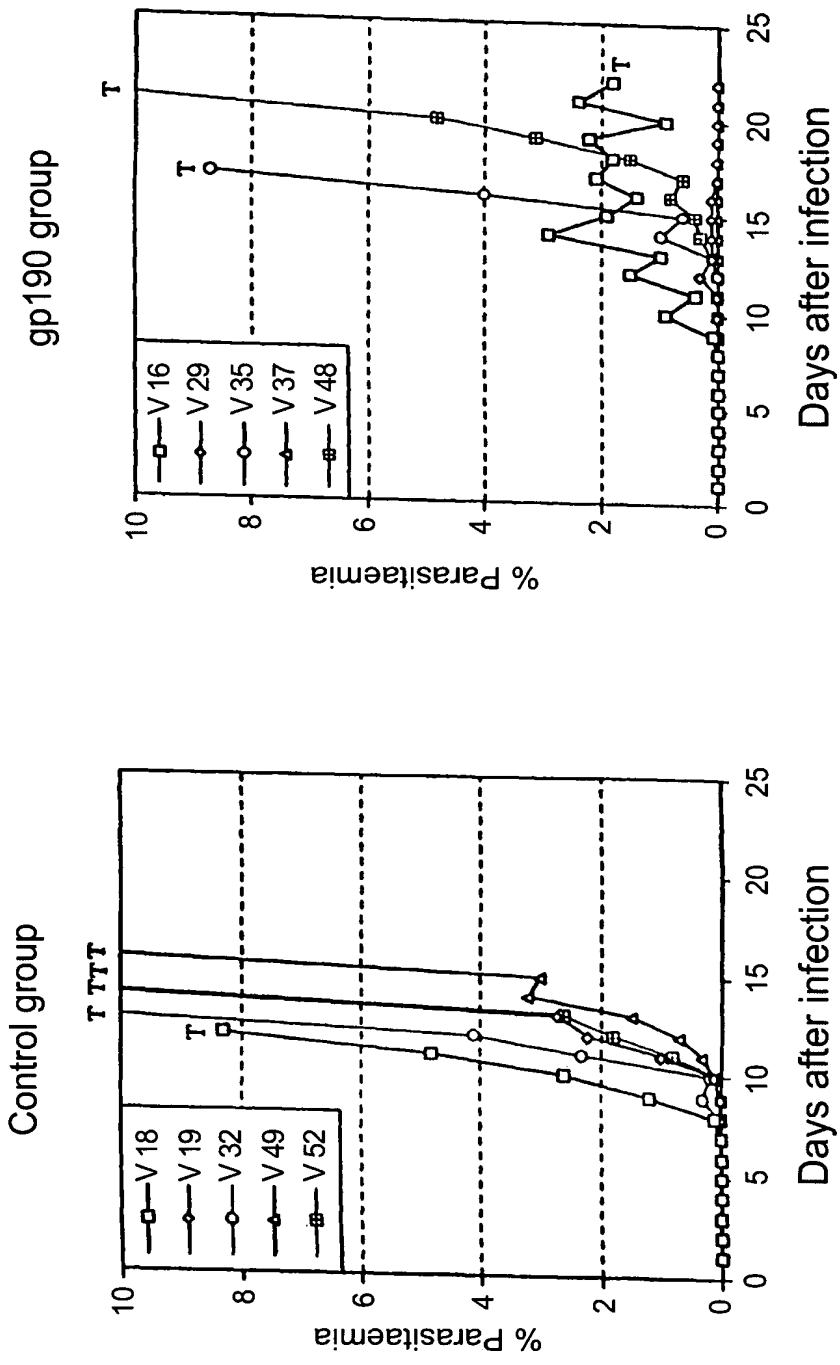
ANNOTATED DRAWING
U.S. Patent Application No. 09/269,874
Edward J. Baba (Reg. No. 52,581)
(650) 833-7731

FIG. 2A



ANNOTATED DRAWING
U.S. Patent Application No. 09/269,874
Edward J. Baba (Reg. No. 52,581)
(650) 833-7731

FIG. 2B



REPLACEMENT DRAWING
U.S. Patent Application No. 09/269,874
Edward J. Baba (Reg. No. 52,581)
(650) 833-7731

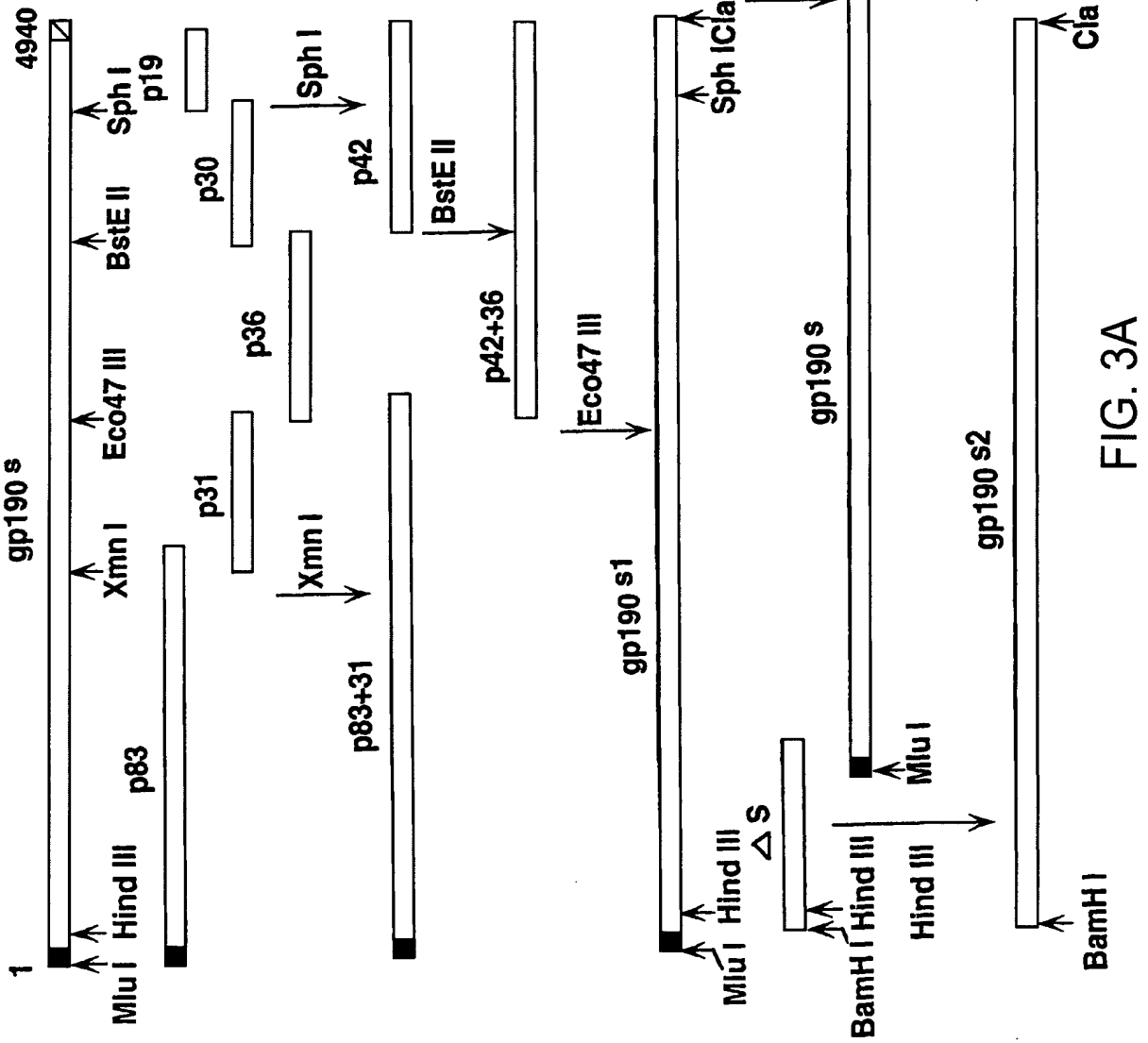


FIG. 3A

REPLACEMENT DRAWING
U.S. Patent Application No. 09/269,874
Edward J. Baba (Reg. No. 52,581)
(650) 833-7731

FIG. 3B

FIG. 3C

DNA sequence of the native (gp190n) and of the synthetic gene (gp190s) for gp190 from FCG-1

```
     AS                M  K  I  I  F  F  L  C  S  F  L  F     12
gp190n                    G     A     TT A           T
gp190s  CGCACGCGTATGAAAATCATTTTCTTCCTCTGTTCATTTCTGTTT     45
           Mlu I

     AS  F  I  I  N  T  Q  C  V  T  H  E  S  Y  Q  E     27
gp190n     T  A     A  A  T  A  A  T   AGT     A  A
gp190s  TTTATCATCAATACTCAGTGCGTGACCCACGAATCCTATCAGGAG     90

     AS  L  V  K  K  L  E  A  L  E  D  A  V  L  T  G     42
gp190n    T  C  A     A        A        A  AT G  A  T
gp190s  CTGGTTAAGAAACTGGAAGCTTTGGAAGATGCCGTCCTTACCGGA    135

     AS  Y  S  L  F  Q  K  E  K  M  V  L  N  E  G  T     57
gp190n    T  TT A  T  A     A  A     AT A        A  A
gp190s  TACAGCCTGTTCCAGAAGGAGAAGATGGTGCTGAATGAAGGGACG    180
```

FIG. 3D

```
    AS  S  G  T  A  V  T  T  S  T  P  G  S  K  G  S    72
gp190n     A  A  T        T  T     T     A  G  T  A
gp190s  AGTGGCACGGCCGTTACAACCAGCACACCCGGTTCTAAAGGGTCT  225

    AS  V  A  S  G  G  S  G  G  S  V  A  S  G  G  S    87
gp190n     T   TCA     T  A     C  A  T  T  A  T  C  A
gp190s  GTGGCTAGCGGTGGCTCCGGTGGGTCTGTGGCCTCTGGGGGTTCC  270

    AS  V  A  S  G  G  S  V  A  S  G  G  S  V  A  S    102
gp190n     T  T  A  T   TCA  T  T                 T  TTCA
gp190s  GTCGCCTCCGGCGGCAGCGTGGCATCAGGTGGCTCAGTGGCAAGC  315

    AS  G  G  S  G  N  S  R  R  T  N  P  S  D  N  S    117
gp190n     T     A  T  TTCAA  C  T  A     T  A  T  T  A
gp190s  GGCGGTTCCGGGAACAGTCGAAGAACCAATCCATCTGACAACTCT  360

    AS  S  D  S  D  A  K  S  Y  A  D  L  K  H  R  V    132
gp190n     T     A  T  T  A  T     T  TT A  A     A
gp190s  AGCGATTCCGACGCCAAGTCCTACGCCGACCTCAAGCACCGAGTG  405

    AS  R  N  Y  L  L  T  I  K  E  L  K  Y  P  Q  L    147
gp190n  C     T  CT GT A        A  A  C  A  T  T  AC C
gp190s  AGAAACTATCTCCTCACTATCAAGGAGCTGAAGTACCCACAGTTG  450
```

FIG. 3E

```
    AS  F  D  L  T  N  H  M  L  T  L  C  D  N  I  H    162
gp190n     T  TT A           T A  TT        T
gp190s  TTCGACCTCACTAATCATATGCTGACACTGTGTGATAACATTCAT  495

    AS  G  F  K  Y  L  I  D  G  Y  E  E  I  N  E  L    177
gp190n     T        T A     T  A  T     A  T     T A
gp190s  GGCTTCAAATATCTGATTGACGGTTACGAAGAGATCAATGAACTC  540

    AS  L  Y  K  L  N  F  Y  F  D  L  L  R  A  K  L    192
gp190n  T A  T  A  A  C  T  T  T  T  AT    A  A  T A
gp190s  CTGTACAAGTTGAATTTCTACTTCGACTTGCTAAGGGCCAAACTG  585

    AS  N  D  V  C  A  N  D  Y  C  Q  I  P  F  N  L    207
gp190n        T  A  T  T     T           A  T     C T
gp190s  AATGACGTTTGCGCCAATGACTATTGTCAAATTCCATTCAATTTG  630

    AS  K  I  R  A  N  E  L  D  V  L  K  K  L  V  F    222
gp190n     A  TC T  A  T  A  A     C T  A  AC T  G
gp190s  AAGATCAGAGCCAACGAGTTGGACGTATTGAAGAAGTTGGTCTTC  675
```

FIG. 3F

```
    AS  G  Y  R  K  P  L  D  N  I  K  D  N  V  G  K    237
gp190n        A  A  A  AT A     T  T  A  T     A     A
gp190s  GGATATCGCAAGCCTCTCGACAACATCAAGGACAATGTGGGAAAG  720

    AS  M  E  D  Y  I  K  K  N  K  K  T  I  E  N  I    252
gp190n              C        A     A  A     A  A  T  A
gp190s  ATGGAAGATTATATTAAAAAGAATAAGAAGACCATCGAGAACATT  765

    AS  N  E  L  I  E  E  S  K  K  T  I  D  K  N  K    267
gp190n    T  AT A  T        AGT  G  A  A  T  T
gp190s  AACGAGCTGATCGAAGAATCCAAAAAGACCATAGACAAAAATAAG  810

    AS  N  A  T  K  E  E  E  K  K  K  L  Y  Q  A  Q    282
gp190n           T  A  A        A  A  A  A     A  T  A
gp190s  AATGCAACCAAGGAGGAAGAAAAGAAGAAGTTGTACCAGGCCCAG  855

    AS  Y  D  L  S  I  Y  N  K  Q  L  E  E  A  H  N    297
gp190n    T  T  T  T  T  C  T     AT A        A     T
gp190s  TACGACCTGTCCATCTATAACAAACAGCTTGAAGAAGCCCATAAC  900
```

FIG. 3G

```
    AS  L  I  S  V  L  E  K  R  I  D  T  L  K  K  N    312
gp190n  T  A  A     TT A  A  A  T  T     TT A  A  A
gp190s  CTCATCAGCGTACTGGAGAAGCGCATAGACACCCTCAAGAAGAAT  945

    AS  E  N  I  K  E  L  L  D  K  I  N  E  I  K  N    327
gp190n     C  T  G  T A  T  T     A              A
gp190s  GAAAATATCAAAGAACTGCTCGACAAGATTAATGAAATTAAGAAT  990

    AS  P  P  P  A  N  S  G  N  T  P  N  T  L  L  D    342
gp190n    C  A  G     T     A  T  A  A  T  T  C  T  T
gp190s  CCTCCGCCAGCCAACTCTGGGAACACCCCTAACACGCTGCTGGAC  1035

    AS  K  N  K  K  I  E  E  H  E  K  E  I  K  E  I    357
gp190n           A  A  C     A     A     A  A     A  T
gp190s  AAGAACAAGAAGATAGAGGAGCACGAGAAAGAGATCAAAGAGATC  1080

    AS  A  K  T  I  K  F  N  I  D  S  L  F  T  D  P    372
gp190n           T     A  T  T  T  AG T A           A
gp190s  GCCAAAACCATTAAGTTCAACATAGATTCTCTCTTTACTGATCCC  1125
```

FIG. 3H

```
    AS  L  E  L  E  Y  Y  L  R  E  K  N  K  N  I  D    387
gp190n     AT A  A     T  A     A  A     A     T  T
gp190s  CTTGAGCTGGAGTACTACTTGAGAGAGAAGAATAAGAATATAGAC  1170

    AS  I  S  A  K  V  E  T  K  E  S  T  E  P  N  E    402
gp190n    AAGT  A  G  T  A              T     C
gp190s  ATCTCCGCCAAAGTCGAGACAAAGGAATCAACCGAACCTAATGAA  1215

    AS  Y  P  N  G  V  T  Y  P  L  S  Y  N  D  I  N    417
gp190n     A     A  T  T  T  T  A              T
gp190s  TATCCCAATGGTGTGACGTACCCTCTGTCTTATAACGATATCAAC  1260

    AS  N  A  L  N  E  L  N  S  F  G  D  L  I  N  P    432
gp190n    T  T A  T  A  T  TCT  T     T  A     T  A
gp190s  AACGCTCTCAACGAGCTCAATAGCTTCGGTGACTTGATTAACCCC  1305

    AS  F  D  Y  T  K  E  P  S  K  N  I  Y  T  D  N    447
gp190n    T        A        AAG   A  C  A  T  T  T
gp190s  TTCGATTATACGAAAGAACCCTCTAAGAATATCTACACAGACAAT  1350
```

FIG. 3I

```
    AS  E  R  K  K  F  I  N  E  I  K  E  K  I  K  I     462
gp190n     A     A  A  C  A  T     T     A  A  T     A
gp190s  GAGAGAAAGAAGTTTATCAACGAAATCAAGGAGAAGATCAAAATT  1395

    AS  E  K  K  K  I  E  S  D  K  K  S  Y  E  D  R     477
gp190n     A  A  A        ATC  T  A  TC           A  A
gp190s  GAGAAGAAGAAAATTGAGAGTGACAAGAAAAGTTACGAAGACCGC  1440

    AS  S  K  S  L  N  D  I  T  K  E  Y  E  K  L  L     492
gp190n  TCT  GTC T     T     A  A     A        AT A  T
gp190s  AGCAAAAGTCTAAACGATATCACTAAAGAGTATGAAAAGCTGCTG  1485

    AS  N  E  I  Y  D  S  K  F  N  N  N  I  D  L  T     507
gp190n     T  A  T     AG        T     T  A  TT A  T
gp190s  AACGAGATCTATGATTCCAAATTCAACAATAACATCGACCTGACC  1530

    AS  N  F  E  K  M  M  G  K  R  Y  S  Y  K  V  E     522
gp190n     T     A           T  A  A  T  A  T     T
gp190s  AACTTCGAGAAAATGATGGGAAAACGGTACTCTTACAAAGTGGAG  1575
```

FIG. 3J

```
    AS  K  L  T  H  H  N  T  F  A  S  Y  E  N  S  K     537
gp190n        T                 T                 A        A
gp190s  AAACTGACACACCATAATACCTTTGCATCCTATGAGAATTCTAAG     1620

    AS  H  N  L  E  K  L  T  K  A  L  K  Y  M  E  D     552
gp190n              A  T  A  A              A        A  T
gp190s  CATAATCTTGAGAAGCTCACCAAAGCTCTTAAGTATATGGAGGAC     1665

    AS  Y  S  L  R  N  I  V  V  E  K  E  L  K  Y  Y     567
gp190n        T  AA    T  A  A  T  A     T     A     T
gp190s  TATTCTCTGCGGAACATTGTTGTGGAGAAAGAACTAAAGTATTAC     1710

    AS  K  N  L  I  S  K  I  E  N  E  I  E  T  L  V     582
gp190n     A     T  A     C  A  A     T     T  A  AT A
gp190s  AAGAATCTCATAAGTAAGATCGAAAACGAGATCGAGACGCTTGTT     1755

    AS  E  N  I  K  K  D  E  E  Q  L  F  E  K  K  I     597
gp190n     A  T     A  A              C  T     A  A  A
gp190s  GAGAACATTAAGAAGGATGAAGAACAGTTGTTTGAGAAGAAGATT     1800
```

FIG. 3K

```
    AS  T  K  D  E  N  K  P  D  E  K  I  L  E  V  S     612
gp190n     T                          A  A  TT A  A  A  T
gp190s  ACAAAAGACGAAAATAAACCAGATGAGAAGATCCTGGAGGTCTCC  1845

    AS  D  I  V  K  V  Q  V  Q  K  V  L  L  M  N  K     627
gp190n     C     A     A     T  A  A  TT AT A        A
gp190s  GATATTGTTAAAGTCCAAGTGCAGAAGGTGCTCCTCATGAACAAG  1890

    AS  I  D  E  L  K  K  T  Q  L  I  L  K  N  V  E     642
gp190n        C  T  A  A        T  G  T  A  A  T  A  A
gp190s  ATTGATGAACTCAAGAAGACTCAACTCATTCTGAAGAACGTGGAG  1935

    AS  L  K  H  N  I  H  V  P  N  S  Y  K  Q  E  N     657
gp190n                     T  C  TC  C  A  A  A
gp190s  TTAAAACATAATATACATGTGCCGAATAGTTATAAGCAGGAGAAT  1980

    AS  K  Q  E  P  Y  Y  L  I  V  L  K  K  E  I  D     672
gp190n        A     T  T  TT A  T  GT G  A     A  T  T
gp190s  AAGCAGGAACCATACTACCTCATCGTACTCAAGAAAGAGATAGAC  2025
```

FIG. 3L

```
    AS  K  L  K  V  F  M  P  K  V  E  S  L  I  N  E    687
gp190n     T A              T  G  A  ATCAT    A  T
gp190s  AAACTGAAAGTGTTCATGCCCAAAGTCGAGAGCCTGATCAACGAA  2070

    AS  E  K  K  N  I  K  T  E  G  Q  S  D  N  S  E    702
gp190n    A  A  A     A     A     T  A  G     T  G  A
gp190s  GAGAAGAAGAACATTAAAACTGAAGGACAGTCAGATAACTCCGAG  2115

    AS  P  S  T  E  G  E  I  T  G  Q  A  T  T  K  P    717
gp190n    A  A  C        A     A     A  A  T  A  A  T
gp190s  CCTTCCACAGAAGGAGAGATAACCGGACAGGCTACCACCAAGCCC  2160

    AS  G  Q  Q  A  G  S  A  L  E  G  D  S  V  Q  A    732
gp190n        A  A  A  T  T A     A  TCA  A     A
gp190s  GGACAACAGGCCGGTTCAGCTCTCGAAGGCGATAGCGTGCAAGCT  2205

    AS  Q  A  Q  E  Q  K  Q  A  Q  P  P  V  P  V  P    747
gp190n           A  A  A  A     A  A     A     A  A
gp190s  CAAGCACAAGAGCAGAAGCAGGCACAGCCTCCAGTGCCAGTGCCC  2250
```

FIG. 3M

```
    AS  V  P  E  A  K  A  Q  V  P  T  P  P  A  P  V    762
gp190n     A     A  A     A     C  A              A  A  A
gp190s  GTTCCAGAGGCTAAAGCTCAAGTGCCTACACCACCAGCTCCTGTG  2295

    AS  N  N  K  T  E  N  V  S  K  L  D  Y  L  E  K    777
gp190n        T  A  T  A     TTC    T  A  T  T     A  A
gp190s  AATAACAAGACCGAGAATGTCAGCAAACTGGACTACCTTGAGAAG  2340

    AS  L  Y  E  F  L  N  T  S  Y  I  C  H  K  Y  I    792
gp190n  T  A     A  TT A     T  A  T  A  T           T
gp190s  CTCTATGAGTTCCTGAATACATCCTACATCTGCCACAAATATATC  2385

    AS  L  V  S  H  S  T  M  N  E  K  I  L  K  Q  Y    807
gp190n  T  G  T  A   TCA           A     AT A     A  T
gp190s  CTCGTCTCTCACAGCACTATGAACGAGAAGATTCTTAAACAGTAC  2430

    AS  K  I  T  K  E  E  E  S  K  L  S  S  C  D  P    822
gp190n     A  T  A     G  A  A  C   T AAGT  A
gp190s  AAGATAACCAAGGAAGAGGAGAGTAAACTGTCCTCTTGTGATCCA  2475
```

FIG. 3N

```
    AS  L  D  L  L  F  N  I  Q  N  N  I  P  V  M  Y    837
gp190n  T  A     T  AT     T        A  A  T        A  T  A
gp190s  CTGGACCTGCTGTTCAATATCCAGAACAACATTCCCGTTATGTAT  2520

    AS  S  M  F  D  S  L  N  N  S  L  S  Q  L  F  M    852
gp190n           T        T  A        AG T  A  A     A  T
gp190s  TCTATGTTCGATAGCCTCAACAATTCTCTCTCTCAACTGTTCATG  2565

    AS  E  I  Y  E  K  E  M  V  C  N  L  Y  K  L  K    867
gp190n     A  T     A  A  A     T  T  TT A           T  G
gp190s  GAGATATATGAGAAGGAGATGGTCTGCAACCTGTATAAACTCAAA  2610

    AS  D  N  D  K  I  K  N  L  L  E  E  A  K  K  V    882
gp190n     T  T     A     A  TT AT A           G  A  A  A
gp190s  GACAACGACAAGATTAAGAACCTTCTGGAGGAAGCTAAGAAGGTC  2655

    AS  S  T  S  V  K  T  L  S  S  S  S  M  Q  P  L    897
gp190n        A     A        T  AAGTTCA  A           T A
gp190s  TCCACCTCTGTTAAAACTCTCTCTTCCAGCTCCATGCAACCACTG  2700
```

FIG. 3O

```
    AS  S  L  T  P  Q  D  K  P  E  V  S  A  N  D  D    912
gp190n    AT A        G  T  A        A  T  A  T  T  T
gp190s  TCTCTCACACCTCAAGACAAGCCCGAAGTGAGCGCTAACGACGAC  2745

    AS  T  S  H  S  T  N  L  N  N  S  L  K  L  F  E    927
gp190n    A  A  T  T  A  TT G     TAGTT A   T A     A
gp190s  ACCTCTCACTCGACCAACCTTAATAACTCACTGAAACTGTTTGAG  2790

    AS  N  I  L  S  L  G  K  N  K  N  I  Y  Q  E  L    942
gp190n       AT  AG   T  A  A  C  A  T  A        T A
gp190s  AACATCCTGTCTCTCGGCAAGAATAAGAACATCTACCAAGAACTT  2835

    AS  I  G  Q  K  S  S  E  N  F  Y  E  K  I  L  K    957
gp190n    A  T  A   AGTAGT  A     T  T  A      T A
gp190s  ATTGGACAGAAATCGTCCGAGAACTTCTACGAGAAGATACTGAAA  2880

    AS  D  S  D  T  F  Y  N  E  S  F  T  N  F  V  K    972
gp190n    T  T  T     T     T  ATCT  T  A  T  T  A
gp190s  GACAGCGACACATTCTATAACGAGAGCTTCACTAACTTCGTGAAA  2925
```

FIG. 3P

```
    AS  S  K  A  D  D  I  N  S  L  N  D  E  S  K  R     987
gp190n           T        T  T  AT G  T        A  A G
gp190s  TCTAAAGCCGATGATATCAACTCTCTTAACGATGAATCTAAACGT  2970

    AS  K  K  L  E  E  D  I  N  K  L  K  K  T  L  Q    1002
gp190n        AT A     A  T  T     AT A  A  A  TT A  G
gp190s  AAGAAGCTGGAAGAGGACATCAATAAGCTGAAGAAGACACTGCAA  3015

    AS  L  S  F  D  L  Y  N  K  Y  K  L  K  L  E  R    1017
gp190n  T ATCA  T  TT A  T  T  A  T   T A   T A  A
gp190s  CTGAGCTTCGACCTGTACAACAAGTACAAACTGAAACTGGAGAGA  3060

    AS  L  F  D  K  K  K  T  V  G  K  Y  K  M  Q  I    1032
gp190n  T A  T  T  A     A     T  T  A     A     A  T
gp190s  CTCTTCGACAAGAAGAAGACAGTCGGCAAGTATAAGATGCAGATC  3105

    AS  K  K  L  T  L  L  K  E  Q  L  E  S  K  L  N    1047
gp190n    A  AC T   T AT A  A  A  AT A   TCA   T G  T
gp190s  AAGAAGTTGACTCTGCTCAAGGAGCAGCTTGAAAGCAAACTCAAC  3150
```

FIG. 3Q

```
    AS  S  L  N  N  P  K  H  V  L  Q  N  F  S  V  F    1062
gp190n     T  T  C  A  G  T  T A  A     T  T  T
gp190s  TCACTGAACAATCCGAAACACGTACTGCAGAACTTCTCAGTGTTC  3195

    AS  F  N  K  K  K  E  A  E  I  A  E  T  E  N  T    1077
gp190n    T     A  A  A     T  A  A  A  A  T  A     A
gp190s  TTCAACAAGAAGAAGGAAGCCGAGATCGCCGAGACAGAGAACACT  3240

    AS  L  E  N  T  K  I  L  L  K  H  Y  K  G  L  V    1092
gp190n  T A  A     A  A  AT AT G     T  T     A  T  T
gp190s  CTGGAGAACACCAAGATTCTTCTCAAACACTACAAAGGCCTCGTC  3285

    AS  K  Y  Y  N  G  E  S  S  P  L  K  T  L  S  E    1107
gp190n    A           T  A  A     AT A  A   T AAGT  A
gp190s  AAGTATTATAATGGCGAGTCTTCTCCTCTGAAGACTCTCTCCGAG  3330

    AS  E  S  I  Q  T  E  D  N  Y  A  S  L  E  N  F    1122
gp190n    ATCA  T  A  A  A     T  T     TT A  A     T
gp190s  GAGAGCATCCAGACCGAGGATAACTACGCCAGCCTCGAGAACTTC  3375
```

FIG. 3R

```
    AS  K  V  L  S  K  L  E  G  K  L  K  D  N  L  N    1137
gp190n    A  AT AAG   AT A     A  AT A     T  TT A  T
gp190s  AAGGTCCTGTCTAAGCTCGAAGGCAAGCTGAAGGACAACCTGAAC  3420

    AS  L  E  K  K  K  L  S  Y  L  S  S  G  L  H  H    1152
gp190n  T A  A     A  AT ATCA  T A  A  T  T A     T
gp190s  CTGGAGAAGAAGAAGCTCAGCTACCTCTCTAGCGGACTGCATCAC  3465

    AS  L  I  A  E  L  K  E  V  I  K  N  K  N  Y  T    1167
gp190n  T A  T  T  AT A  A     A  A  A  T  A  T  T  A
gp190s  CTGATCGCCGAGCTCAAGGAAGTCATTAAGAACAAGAACTACACC  3510

    AS  G  N  S  P  S  E  N  N  T  D  V  N  N  A  L    1182
gp190n    T   TCT     T  A        G  T  T  C  T   T A
gp190s  GGCAATAGCCCAAGCGAGAATAATACAGACGTGAATAACGCACTG  3555

    AS  E  S  Y  K  K  F  L  P  E  G  T  D  V  A  T    1197
gp190n             A  A  T  C  A              T  A  A
gp190s  GAATCTTACAAGAAGTTCCTGCCTGAAGGAACAGATGTCGCCACT  3600
```

FIG. 3S

```
    AS  V  V  S  E  S  G  S  D  T  L  E  Q  S  Q  P    1212
gp190n     T  AAG     AG    A           T A  A  AAG       A
gp190s  GTGGTGTCTGAATCTGGCTCCGACACACTGGAGCAGTCTCAACCT   3645

    AS  K  K  P  A  S  T  H  V  G  A  E  S  N  T  I    1227
gp190n        A  A     A        A     A     T  C     A
gp190s  AAGAAGCCTGCATCTACTCATGTCGGAGCCGAGTCCAATACAATT   3690

    AS  T  T  S  Q  N  V  D  D  E  V  D  D  V  I  I    1242
gp190n     A     A  A  T     T     A  A        A     A
gp190s  ACCACATCTCAGAACGTCGACGATGAGGTCGATGACGTCATCATT   3735

    AS  V  P  I  F  G  E  S  E  E  D  Y  D  D  L  G    1257
gp190n     A     A  T  A  ATC   A  A  T  T     TT A  A
gp190s  GTGCCTATCTTCGGCGAGAGCGAGGAGGACTACGATGACCTCGGC   3780

    AS  Q  V  V  T  G  E  A  V  T  P  S  V  I  D  N    1272
gp190n     A  A  A  A  A  A  A  A           A
gp190s  CAGGTGGTCACCGGTGAGGCTGTCACTCCTTCCGTGATTGATAAC   3825
```

FIG. 3T

```
    AS  I  L  S  K  I  E  N  E  Y  E  V  L  Y  L  K    1287
gp190n     A  T  T     T  A  T     T  G  TT A   T A
gp190s  ATTCTGTCCAAAATCGAGAACGAATACGAAGTGCTCTATCTGAAA  3870

    AS  P  L  A  G  V  Y  R  S  L  K  K  Q  L  E  N    1302
gp190n     T A     T  T     AAG T A  A     AT A  A
gp190s  CCTCTGGCAGGCGTCTATAGGTCTCTCAAGAAACAGCTGGAGAAT  3915

    AS  N  V  M  T  F  N  V  N  V  K  D  I  L  N  S    1317
gp190n        T     A  T     T  T  T     T   T A  TTCA
gp190s  AACGTGATGACCTTCAATGTCAACGTGAAGGACATTCTGAACAGC  3960

    AS  R  F  N  K  R  E  N  F  K  N  V  L  E  S  D    1332
gp190n     A        AC T           A  T  T  A  ATCA  T
gp190s  CGCTTTAATAAGAGAGAAAATTTCAAGAACGTCTTGGAGAGCGAC  4005

    AS  L  I  P  Y  K  D  L  T  S  S  N  Y  V  V  K    1347
gp190n     A     A        TT A  A  AAG   T  T        A
gp190s  TTGATTCCCTATAAAGACCTGACCTCCTCTAACTACGTTGTCAAG  4050
```

FIG. 3U

```
    AS  D  P  Y  K  F  L  N  K  E  K  R  D  K  F  L    1362
gp190n     T     T  A  T  T        A  A  A        CT A
gp190s  GACCCATACAAGTTCCTCAATAAAGAGAAGAGGGATAAATTTCTG  4095

    AS  S  S  Y  N  Y  I  K  D  S  I  D  T  D  I  N    1377
gp190n  AGC     T  T     T     T  A  A  T  G     A
gp190s  TCTAGTTACAACTATATCAAGGACTCCATCGACACCGATATCAAT  4140

    AS  F  A  N  D  V  L  G  Y  Y  K  I  L  S  E  K    1392
gp190n     T  A        T  T  A     T  A  AT ATC
gp190s  TTCGCTAATGATGTGCTGGGGTATTACAAGATCCTGAGCGAAAAA  4185

    AS  Y  K  S  D  L  D  S  I  K  K  Y  I  N  D  K    1407
gp190n     T  A  A  TT A  T  A        A           C  A
gp190s  TACAAGTCTGACCTTGACTCTATTAAAAAGTATATCAACGATAAG  4230

    AS  Q  G  E  N  E  K  Y  L  P  F  L  N  N  I  E    1422
gp190n        T  A     G     C  T     TT A  C  T  T  G
gp190s  CAAGGCGAGAATGAAAAATATCTGCCCTTCCTGAATAACATCGAA  4275
```

FIG. 3V

```
    AS  T  L  Y  K  T  V  N  D  K  I  D  L  F  V  I    1437
gp190n     T  A  T  A     T  T  T     T  TT A  T
gp190s  ACCCTGTACAAGACAGTGAACGACAAAATCGACCTCTTCGTAATT   4320

    AS  H  L  E  A  K  V  L  N  Y  T  Y  E  K  S  N    1452
gp190n    TT A  A  A  A  T  A  T     A  T     ATCA  C
gp190s  CACCTGGAGGCCAAGGTCCTCAACTATACTTACGAGAAGAGCAAT   4365

    AS  V  E  V  K  I  K  E  L  N  Y  L  K  T  I  Q    1467
gp190n    A              A  A  A  T  T  T A        T
gp190s  GTGGAAGTTAAAATCAAGGAGCTGAACTACCTCAAAACAATCCAA   4410

    AS  D  K  L  A  D  F  K  K  N  N  N  F  V  G  I    1482
gp190n        AT           T  A                    T
gp190s  GACAAGCTGGCAGATTTCAAGAAAAATAACAATTTCGTCGGAATT   4455

    AS  A  D  L  S  T  D  Y  N  H  N  N  L  L  T  K    1497
gp190n    T  TT A  A  A           T  T  CT AT    A
gp190s  GCAGACCTGTCTACCGATTATAACCACAACAATCTCCTGACCAAG   4500
```

FIG. 3W

```
    AS  F  L  S  T  G  M  V  F  E  N  L  A  K  T  V    1512
gp190n    C  TAGT  A  T     T  T     T  T  T     C  T
gp190s  TTTCTGTCCACTGGCATGGTGTTCGAAAACCTCGCCAAAACAGTG  4545

    AS  L  S  N  L  L  D  G  N  L  Q  G  M  L  N  I    1527
gp190n  T ATCT   T A  T  T  A  T    A  T   T A     T
gp190s  CTGAGCAATCTGCTCGACGGCAACCTGCAGGGCATGCTGAACATC  4590

    AS  S  Q  H  Q  C  V  K  K  Q  C  P  Q  N  S  G    1542
gp190n    A  A           A  A     A  T  A  A   TCT  A
gp190s  TCCCAGCACCAATGCGTGAAGAAACAGTGCCCCCAGAATAGCGGC  4635

    AS  C  F  R  H  L  D  E  R  E  E  C  K  C  L  L    1557
gp190n        A   T A  T  AA A    A  T  A   T AT A
gp190s  TGTTTCAGGCATCTGGACGAGCGCGAAGAGTGCAAGTGTCTCCTG  4680

    AS  N  Y  K  Q  E  G  D  K  C  V  E  N  P  N  P    1572
gp190n    T              T     A  T  T  A  T     T
gp190s  AACTACAAACAAGAAGGAGATAAGTGCGTGGAGAACCCAAACCCT  4725
```

FIG. 3X

```
    AS  T  C  N  E  N  N  G  G  C  D  A  D  A  K  C    1587
gp190n     T  T  C     T     T  A     T  A     C     T
gp190s  ACCTGCAATGAAAACAATGGCGGGTGTGACGCCGATGCTAAATGC  4770

    AS  T  E  E  D  S  G  S  N  G  K  K  I  T  C  E    1602
gp190n        A     TTCA  TAGC                    T  A
gp190s  ACCGAGGAAGACAGCGGCTCTAACGGAAAGAAAATCACATGCGAG  4815

    AS  C  T  K  P  D  S  Y  P  L  F  D  G  I  F  C    1617
gp190n           A  T  T  T        T     T  T     C
gp190s  TGTACTAAGCCCGACTCCTATCCACTCTTCGACGGGATTTTTTGC  4860

    AS  S  S  S  N  F  L  G  I  F  F  L  L  I  L  M    1632
gp190n  AGTTC     C  T  A  A  A CA  T AT A  A
gp190s  TCCAGCTCTAATTTCCTGGGCATCTTCTTCCTGCTGATCCTCATG  4905

    AS  L  I  L  Y  S  F  I  *  *              1639
gp190n  T  A  AT A     T     T
gp190s  CTGATCCTGTACAGCTTCATCTAATAGATCGATGG     4940
                            stop codon  Cla I
```

FIG. 3Y

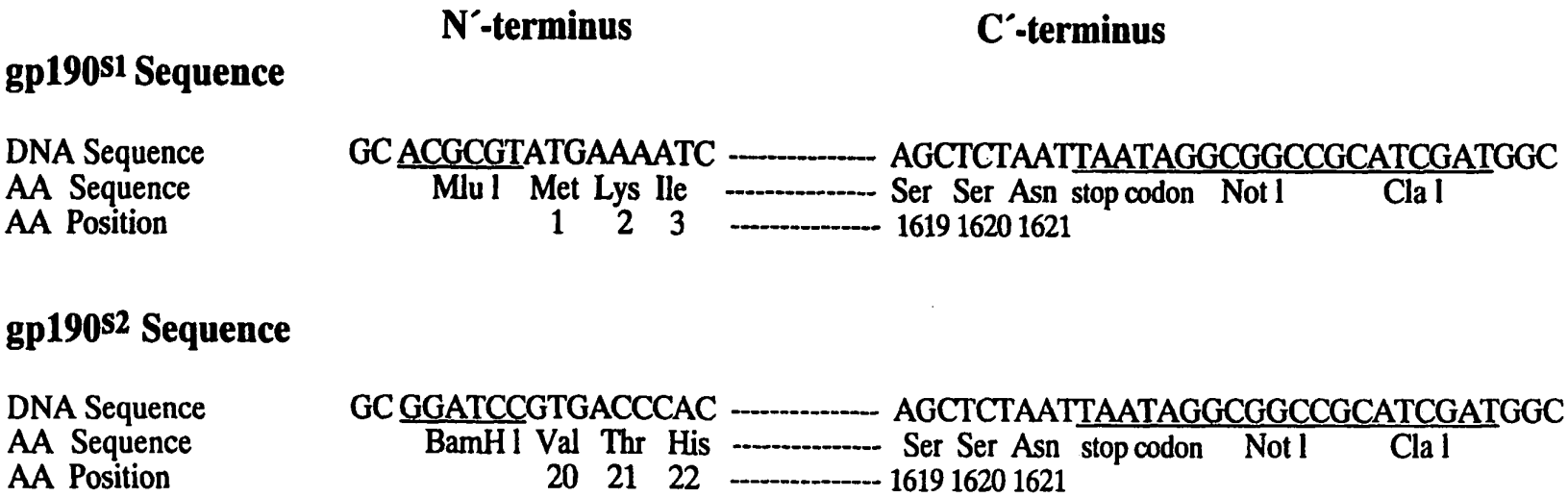
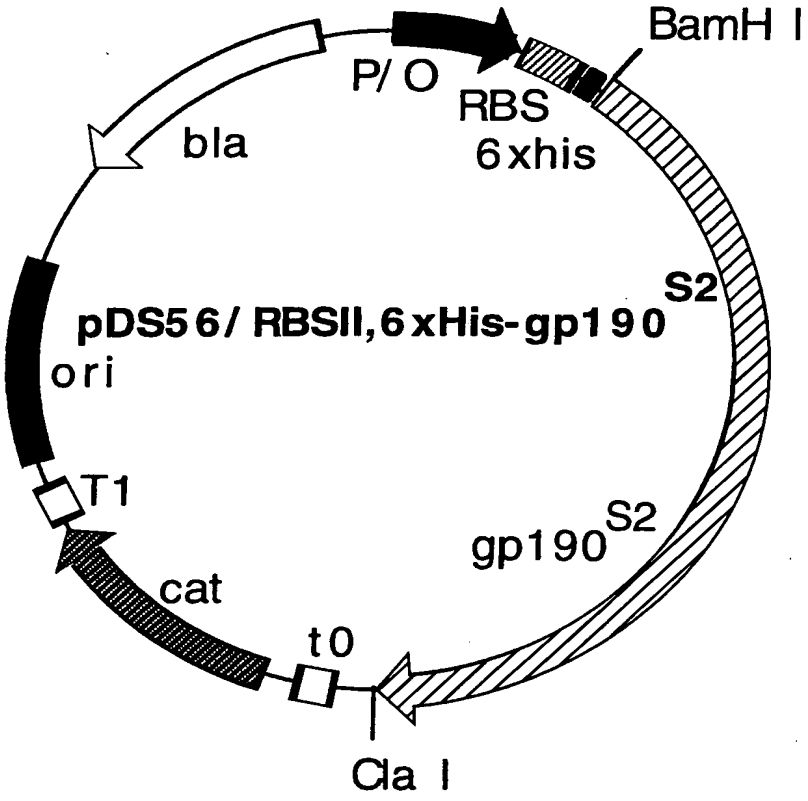
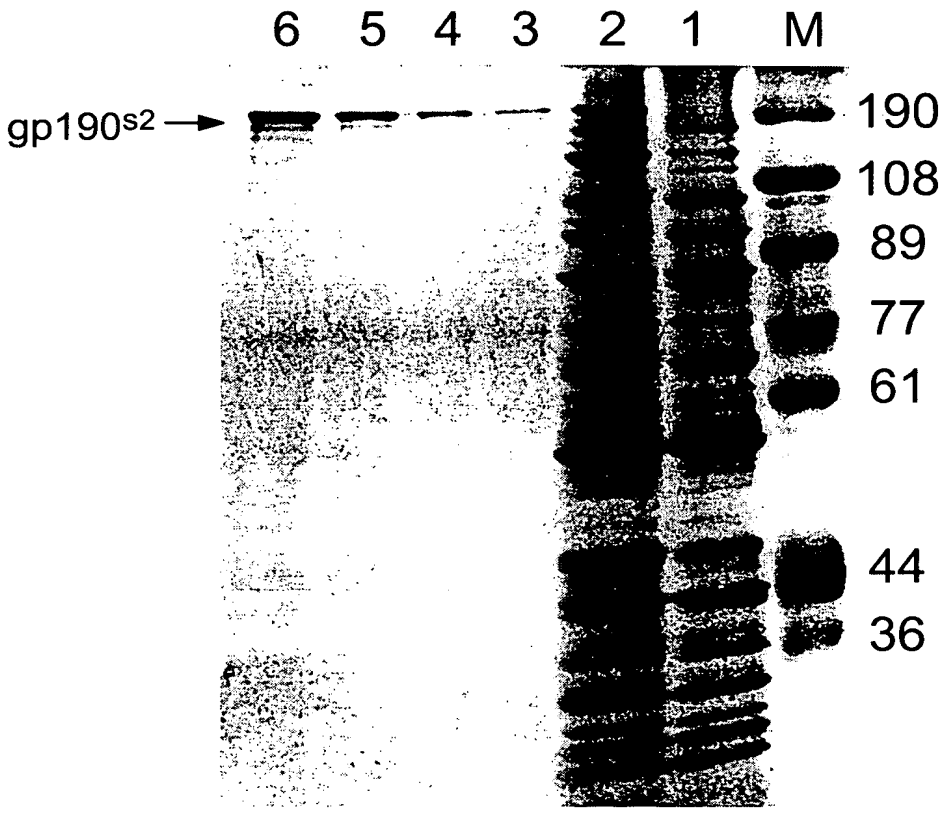


FIG. 4A



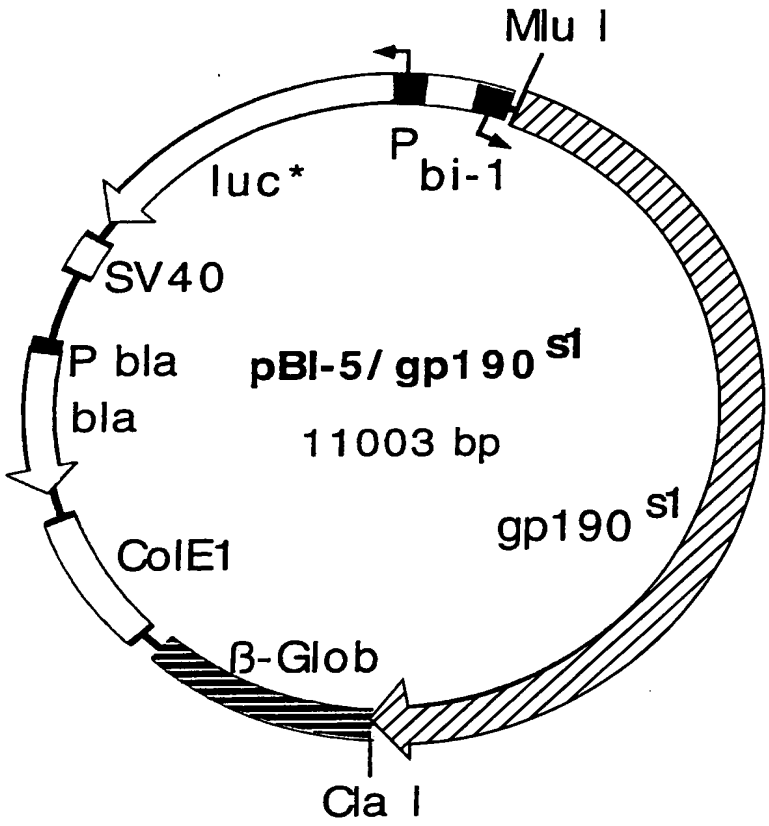
REPLACEMENT DRAWING
U.S. Patent Application No. 09/269,874
Edward J. Baba (Reg. No. 52,581)
(650) 833-7731

FIG. 4B



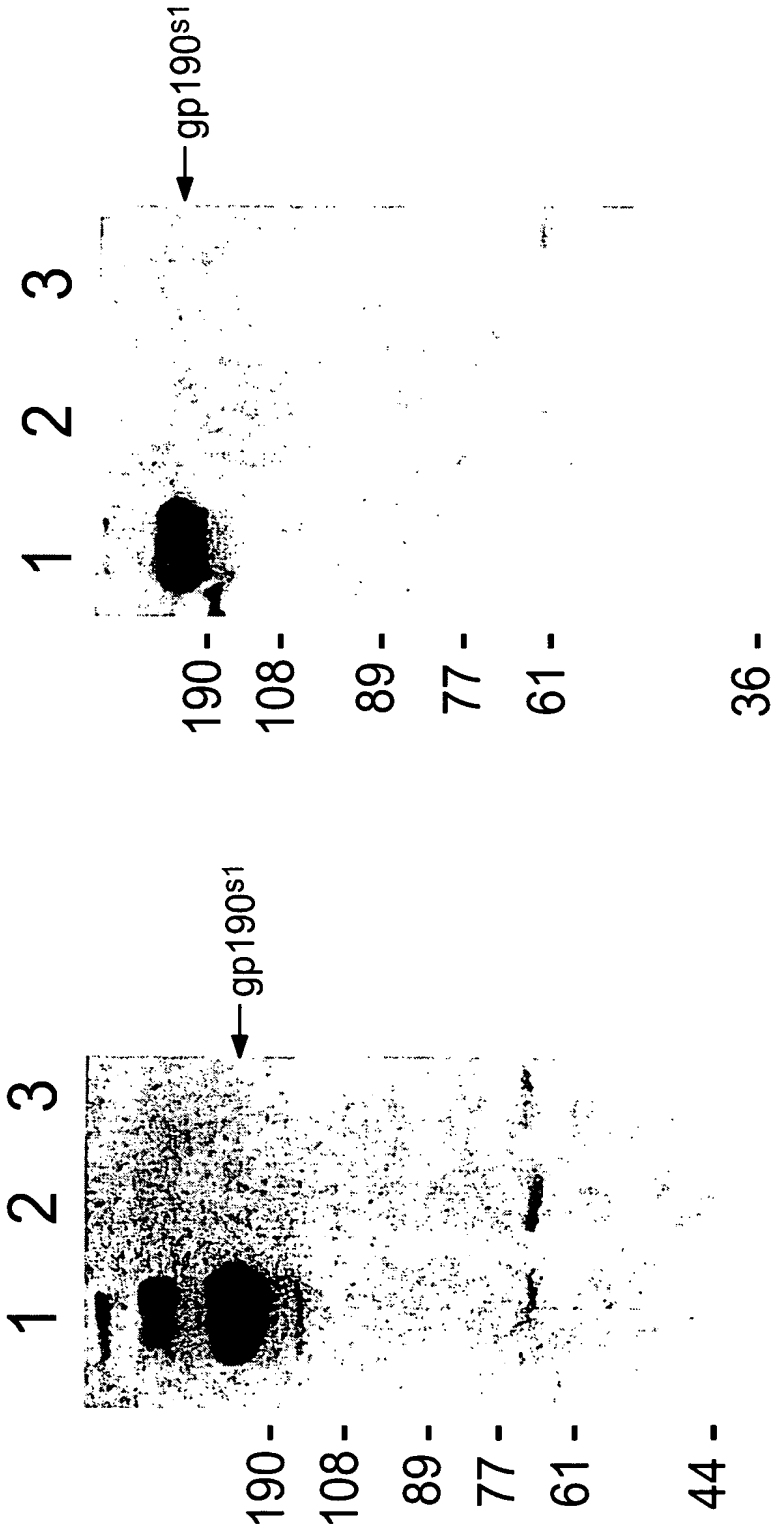
REPLACEMENT DRAWING
U.S. Patent Application No. 09/269,874
Edward J. Baba (Reg. No. 52,581)
(650) 833-7731

FIG. 5A



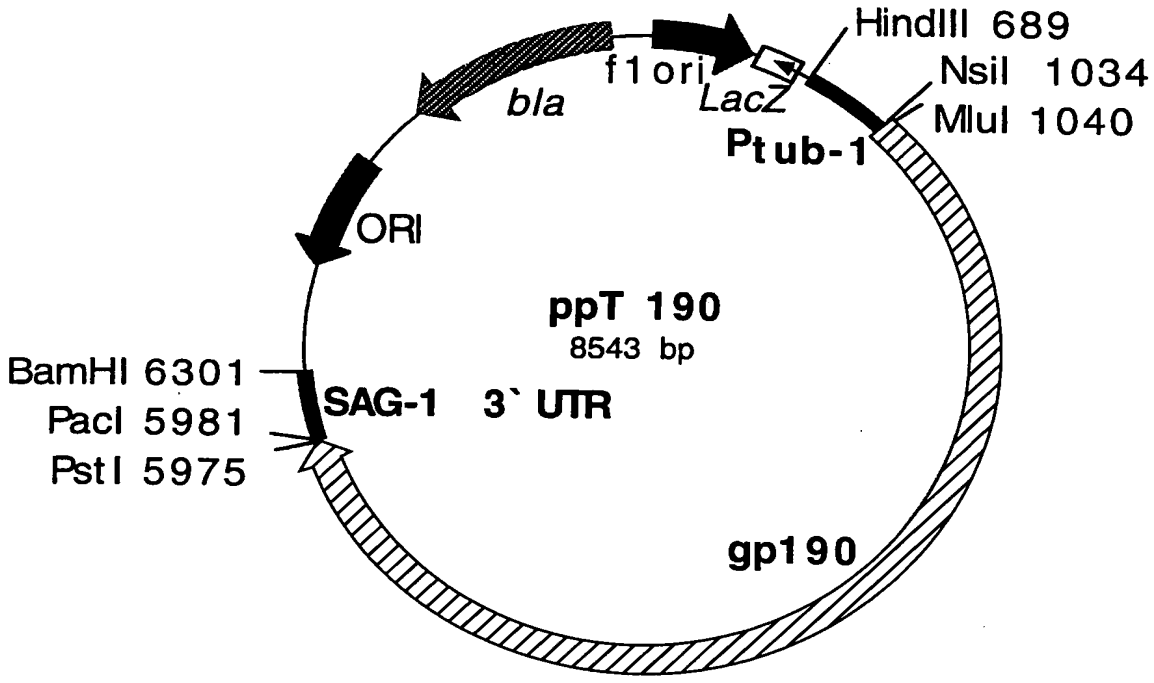
ANNOTATED DRAWING
U.S. Patent Application No. 09/269,874
Edward J. Baba (Reg. No. 52,581)
(650) 833-7731

FIG. 5B



REPLACEMENT DRAWING
U.S. Patent Application No. 09/269,874
Edward J. Baba (Reg. No. 52,581)
(650) 833-7731

FIG. 6A



ANNOTATED DRAWING
U.S. Patent Application No. 09/269,874
Edward J. Baba (Reg. No. 52,581)
(650) 833-7731

FIG. 6B